

Delaval, Jan

110634

From: Roark, Jessica
Sent: Wednesday, December 17, 2003 2:44 PM
To: Delaval, Jan
Subject: 10/045180

Jan,

When things are up and running.....

Please search, including pending, the following from 10/045180:

SEQ ID NO:3 against the DNA database
SEQ ID NO:3 as an oligo against protein

SEQ ID NO:6 against DNA
SEQ ID NO:6 as an oligo against protein

SEQ ID NO:4 against DNA
SEQ ID NO:5 against DNA

Results on paper please.

Thanks!

Jessica H. Roark

CM1 8A03
Mailbox 9E12
Art Unit 1644
703 605-1209

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 2689.5 Seconds
(without alignments)
1429.821 Million cell updates/sec

Title: US-10-045-180A-3
Perfect score: 514
Sequence: 1 MRTLTLASFLVALQAAE.....GEHLGGTCFILGERYPICCY 94

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool.p/US10045180/runat.17122003.145023.12528/app.query.fasta.1.860
-DB=GenEmbl -QFWT=fastap -SUPFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180 @CGN.1.1.5283 @runat.17122003.145023.12528 -NCFU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORE=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_on:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*

RESULT 1

ALIGNMENTS

Result No.	Score	Query Match	Length	DB ID	Description
1	514	100.0	453	6	A98571 Sequence 2
2	514	100.0	453	6	BD074746 Human def
3	393.5	76.6	4415	6	A98570 Sequence 1
4	393.5	76.6	4415	6	BD074745 Human def
5	374	72.8	128544	9	AF238378 Homo sapi
6	300	58.4	163093	2	AC116559 Papio anu
7	300	58.4	188604	2	AC116558 Papio anu
8	227	44.2	542	9	HSCC8HP4
9	225.5	43.9	347	9	AF188268 Macaca mu
10	225.5	43.9	435	9	AF184160 Macaca mu
11	223.5	43.5	433	9	AF184159 Macaca mu
12	220	42.8	452	6	AR270635 Sequence
13	220	42.8	452	6	I49588 Sequence 6
14	220	42.8	452	9	HUMDFNSIX
15	218.5	42.5	448	9	HUMDFE1A
16	218.5	42.5	478	6	BD028213 Sequence
17	218.5	42.5	482	9	HSHP1
18	218.5	42.5	498	6	BD174049
19	218.5	42.5	498	9	HUMDFE1A
20	218.5	42.5	514	6	AX405718 Sequence
21	217.5	42.3	347	9	AF188270 Macaca mu
22	216.5	42.1	451	9	HUMDFE1A
23	216.5	42.1	464	9	HSDEF3
24	216.5	42.1	556	9	BC027917 Homo sapi
25	211.5	41.1	347	9	AF188269 Macaca mu
26	207	40.3	424	6	I49587 Sequence 4
27	202	38.3	485	4	RABDNP4A
28	200	38.9	495	4	RABMCP1A
29	199	38.7	479	4	RABNPCS4A
30	199	38.7	500	4	S55582
31	199	38.7	532	4	RABMCP2A
32	198	38.5	454	4	RABNPCS6A
33	198	38.5	464	4	S55578
34	198	38.5	497	4	RABDNP5A
35	197.5	38.4	543	6	AX588765
36	196.5	38.2	469	4	RABDNP3AB
37	195	37.9	464	10	RNU16685
38	194	37.7	485	10	RNU16686
39	186	36.2	504	10	RNU16683
40	175.5	34.1	496	9	AF355799
41	173.5	33.8	420	6	AR059558
42	173.5	33.8	420	6	AR063040 Sequence
43	173.5	33.8	420	6	I93525 Sequence 46
44	171.5	33.4	401	6	AR059566 Sequence
45	171.5	33.4	401	6	AR063048 Sequence

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

FEATURES	source	Location/Qualifiers
		1..4415
		/organism="unidentified"
		/mol_type="genomic DNA"
		/db_xref="taxon:32644"
exon		1836..1874
exon		3394..3577
exon		4161..4380
polyA site		4374..4379
BASE COUNT	1128 a 1109 c 912 g 1252 t	14 others
ORIGIN		
Alignment Scores:		
Pred. No.:	5.07e-39	Length: 4415
Score:	393.50	Matches: 94
Percent Similarity:	32.41%	Conservative: 0
Best Local Similarity:	32.41%	Mismatches: 0
Query Match:	76.56%	Indels: 196
DB:	6	Gaps: 1
US-10-045-180A-3 (1-94) x A98570 (1-4415)		
Qy	1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTtpAlaGlu	20
Db	3406 ATGAGGACCTCACCCCTCTCTGCTCTCTCTCTGCTGCGCCCTTACGGCTCTGGCAGAG	3465
Qy	21 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAspGln	40
Db	3466 CCGCTCCAGCAGAGAGCTCATGAGATGCCAGGCCAGAGCAGCCTCCAGCAGATGACCAG	3525
Qy	41 AspValValIleTyrPheSerGlyAspSerCysSerLeuGln	55
Db	3526 GATGTGTCATTTACTTTTTCAGGAGATGACAGCTGCTCTCTTCAGGTTCAGGTGAGAGA	3585
Qy	55 -----	55
Db	3586 TGCCAGCATGCAGAGCTACAGACTAGACAGAGGACAGGAGACAGGCTCTGGAAATGGAT	3645
Qy	55 -----	55
Db	3646 CTCAGTGCAGATGTCACCTTAGTGGCTATACTTAACATCTCTGGTCTCGATTTTCTCA	3705
Qy	55 -----	55
Db	3706 TATCTAAATGGAATAGAGAACCAAGAAATCTAAGAGATTTTCTTCTCCAAAAAATTG	3765
Qy	55 -----	55
Db	3766 ATTCCAAGATATGACTGTGAAATTCATAGATTTAAGATATAAGGAGATGCTACCTAGTT	3825
Qy	55 -----	55
Db	3826 CCTTCTGGAGCCAGACAAACAAGCTTAAGTATATAGGAAAATATTTCACCTGCTATAT	3885
Qy	55 -----	55
Db	3886 AGGAGGTTTTAGAACCTGGAGAGGCGCTAAGAATGTGTTTCAGGTGTGTGTGATGGGC	3945
Qy	55 -----	55
Db	3946 AGGAATGCAGAAAAAGTGAACAAAGGAGAAATGAGTCTCGAATCCTGTGTGACCGACACTG	4005
Qy	55 -----	55
Db	4006 CTCGTGTATTTATTTCTATTGACTGAGATGTTTGTGTTACCGCTGTAAATACAGCAA	4065
Qy	55 -----	55
Db	4066 CATCACTCATCAGCAACATGTGACTTCTCCAAGATTTCCCTTTTACCACCACCTGCTGNAC	4125
Qy	56 -----	65
Db	4126 CCCGTACTCAGTTTCTGATGCTCTCTCTGGGTCCCGAGGTCAACAAAGGCTTGATCTG	4185


```
QY 21 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspGln 40
Db 3466 CCGCTCCAGCAGAGCTCATGAGATGGCCAGCCAGAGCAGCTCCAGCAGATGACCA 3525
QY 41 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGln----- 55
Db 3526 GATGGTCAATTACTTTTCAGGAGATGACAGCTGCTCTTCAGGTTCCAGGTGAGAGA 3585
QY 55 ----- 55
Db 3586 TGCCAGCATGCAGAGCTACAGACTAGACAGAGGACAGGAGCAGGCTCTGGAATTGGAT 3645
QY 55 ----- 55
Db 3646 CTCAGTGGCAGATGTCACCTAGGTGGCTATACTTAACATCTCTGGTCCTGGATTTCTCA 3705
QY 55 ----- 55
Db 3706 TATCTAAATGGAATAGAGAACCAAGAAATCTAAGAGATTTTCTTCTCCAAAACTTG 3765
QY 55 ----- 55
Db 3766 ATTCCAAGATATCACTGTGAAATTCACATAGATTAAAGATATAAGGAGATGCTACCTAGTT 3825
QY 55 ----- 55
Db 3826 CTTCTGGAGCCAGACAAACAGCTTAAGTATATATAGGAAATATTTCAACCTGCTATAT 3885
QY 55 ----- 55
Db 3886 AGGAGGTTTTAGAACCTGGAGAGGAGCCCTAAGAATGTGTTTCAGGTGTGTGTGATGGC 3945
QY 55 ----- 55
Db 3946 AGGAATGCAGAAAAGTGAAGCAAGAGAGATGAGTCTCGAATCCTGTGTGACGACACTG 4005
QY 55 ----- 55
Db 4006 CTCTGTGTATTATTTCTATTGACTGAGATGTTGTGTGTGTGTGTGTGTGTGTGTGTGT 4065
QY 55 ----- 55
Db 4066 CATCACTCATCAGCAACATGTGACTTCTCCAAGATTCCTTTACCAACCCACTGCTGNAC 4125
QY 56 -----Val-ProGlySerThrIysGlyLeuIleCys 65
Db 4126 CCGGTACTCATGTTCTGATGCTCTCTGTGGTCCCGAGGCTCAACAAAGGGCTTGATCTG 4185
QY 65 SHIScysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIleLe 85
Db 4186 CCATTGCAGAGTACTATPACTGCAATTTTGGAGAACATCTTGGTGGGACCTGCTTCATCCT 4245
QY 85 uGlyGluArgTyrProIleCysCysTyr 94
Db 4246 TGGTGAACGTTACCCATCTGCTGCTAC 4273

RESULT 5
AF238378/c
LOCUS
DEFINITION Homo sapiens chromosome 8 clone SCB-561b17 map p22-p21, complete
sequence.
ACCESSION AF238378
VERSION AF238378.5 GI:29469504
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 128544)
Polley,A., Baumgart,C., Blechschmidt,K., Dette,M.D., Jahn,N.,
Menzel,U., Reichwald,K., Schilhabel,M.B., Schudy,A., Taudien,S.,
Wen,G., Schutte,B., Malik,M., Peng,J.Hong., McCray,P. and
Rosenthal,A.
```

```
Chromosome 8 genomic sequence
Unpublished
2 (bases 1 to 128544)
Schudy,A., Schilhabel,M., Schutte,B., Ganz,T., Linzmeier,R.,
Ho,C.H., Hoang,B.V., McCray,P., Baumgart,C., Menzel,U.,
Schattevoy,R. and Rosenthal,A.
Direct Submission
Submitted (22-FEB-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
3 (bases 1 to 128544)
Genome Sequencing Center Jena.
Direct Submission
Submitted (24-MAY-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
4 (bases 1 to 128544)
Genome Sequencing Center Jena.
Direct Submission
Submitted (10-NOV-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
5 (bases 1 to 128544)
Lagemann,D. and Platzer,M.
Direct Submission
Submitted (06-JUL-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
6 (bases 1 to 128544)
Lagemann,D. and Platzer,M.
Direct Submission
Submitted (02-APR-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On Apr 2, 2003 this sequence version replaced gi:21700555.
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H370
Center clone name: SCB-561b17
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 127999 bases at least Q40
Consensus quality: 128330 bases at least Q30
Consensus quality: 128544 bases at least Q20
Quality coverage: 11.05x
-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest.
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.
-----
FEATURES
source
1..128544
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="p22-p21"
/clone="SCB-561b17"
```

```
misc_feature 44529..44595 /note="single stranded/single chemistry region"
misc_feature 45703 /note="low quality region"
misc_feature 45815..45872 /note="low quality region"
misc_feature 45824..45872 /note="single stranded/single chemistry region"
unsure 45843
unsure 45847
unsure 45849
unsure 45858
unsure 45868
misc_feature 46235 /note="low quality region"
misc_feature 47899..47900 /note="low quality region"
unsure 47899..47900
misc_feature 49172 /note="low quality region"
misc_feature 49172
misc_feature 49339 /note="low quality region"
misc_feature 49756..49807 /note="single stranded/single chemistry region"
misc_feature 49961 /note="low quality region"
misc_feature 52933..52948 /note="low quality region"
misc_feature 52943 /note="low quality region"
misc_feature 53544 /note="low quality region"
misc_feature 53612..53926 /note="single stranded/single chemistry region"
misc_feature 53674..53677 /note="single chemistry region"
misc_feature 53779..53799 /note="low quality region"
unsure 53787
misc_feature 54734..55129 /note="single stranded/single chemistry region"
misc_feature 56028..56165 /note="single stranded/single chemistry region"
misc_feature 57952..58060 /note="single stranded/single chemistry region"
misc_feature 57968..57974 /note="low quality region"
unsure 57968
unsure 57970
unsure 57974
misc_feature 58652 /note="low quality region"
misc_feature 59561..59638 /note="single stranded/single chemistry region"
misc_feature 59925 /note="low quality region"
unsure 59925
misc_feature 59929 /note="low quality region"
misc_feature 61695..62103 /note="single stranded/single chemistry region"
misc_feature 61858..61859 /note="low quality region"
misc_feature 62263..62485 /note="single stranded/single chemistry region"
misc_feature 62655..63118 /note="single stranded/single chemistry region"
misc_feature 62750..62785 /note="single stranded/single chemistry region"
unsure 62774..62777 /note="low quality region"
unsure 62783
```

```
misc_feature 62801..62825 /note="low quality region"
unsure 62819
unsure 62822
misc_feature 62872..62874 /note="low quality region"
misc_feature 62908 /note="low quality region"
misc_feature 62917..62987 /note="low quality region"
unsure 62933..62935
unsure 62955
unsure 62957
misc_feature 63082 /note="low quality region"
misc_feature 63110 /note="low quality region"
misc_feature 63119..63281 /note="low quality region"
misc_feature 63119..63141 /note="single clone coverage"
unsure 63129..63131
misc_feature 63142..63258 /note="single stranded/single chemistry region"
unsure 63159
unsure 63197..63258
unsure 63281
misc_feature 63876..63878 /note="low quality region"
misc_feature 63933..64021 /note="single stranded/single chemistry region"
misc_feature 63943 /note="low quality region"
misc_feature 63984..63985 /note="low quality region"
misc_feature 64004..64007 /note="low quality region"
misc_feature 70360 /note="low quality region"
misc_feature 70360 /note="low quality region"

Alignment Scores:
Pred. No.: 6 61e-35 Length: 128544
Score: 374.00 Matches: 90
Percent Similarity: 31.60% Conservative: 1
Best Local Similarity: 31.25% Mismatches: 2
Query Match: 72.76% Indels: 196
DB: 9 Gaps: 1

US-10-045-180A-3 (1-94) x AF238378 (1-128544)
QY 2 ArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGluPro 21
Db 63320 AGGACCCCTCACCCCTCTCTCTGCCCTTCTCTGGTGGCCCTTCAGGCCCTGGCAGAGCCG 63261
QY 22 LeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProAlaHisPaspGlnAsp 41
Db 63260 CTCGCGCAAGAGCTCATGAGATGCCAGCCAGAGCAGCCCTCCAGCAGATGACCAGAT 63201
QY 42 ValValIleTyPheSerGlyAspAspSerCysSerLeuGlnVal----- 56
Db 63200 GGGTCCCTTACTTTTCAGGAGATGACAGCTCTCTCTTCAGGT-TCCAGGTGAGAGATG 63142
QY 56 ----- 56
Db 63141 CCAGCATGCAGAGCTACAGACTAGACAGAGCAGAGCAGGCTCTGGAATTGGATCT 63082
QY 56 ----- 56
Db 63081 CAGTGCAGATGTCACCTAGGTGGCTATACTTAAACATCTCTGGTCTGGATTTTCTCAT 63022
QY 56 ----- 56
Db 63021 TCTAAATGGAATAGAGAACCAAGAAATCTAAGAGATTTTCTTTCTCCAAAACTTGAT 62962
```



```
QY 56 ----- 56
Db 33475 TATTCTATTGACTGAGATTGTTCTGCTACGGCTGCAATACAGCCAACTCACTCATC 33534

QY 56 ----- 56
Db 33535 AGTCACAGGTGACTTCTCCAGATTCTTACCATCCACTCGGACCCCATACTGAAT 33594

QY 57 -----
Db 33595 TTCTGATGCTCTCTCTGTTCTCCAGGCTCAACGAAGGGCTTGAGCTGCCATTGCAGAGT 33654

QY 69 lleuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIleLeuGlyGluArgTy 89
Db 33655 ACTTACTAGCCTTTTGTGAGAACATCTTGGTGGACCTCTTCATCCATGGTGAACGCTC 33714

QY 89 rProIleCysCysTyr 94
Db 33715 CCCAATCTGCTGCTAC 33730

RESULT 7
AC116558/c
LOCUS 188604 bp DNA linear HTG 11-JUN-2003
DEFINITION Papio anubis clone rp41-273g19, WORKING DRAFT SEQUENCE, 2 ordered
          pieces.
ACCESSION AC116558
VERSION 3
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
          Cercopitheciinae; Papio.
REFERENCE 1 (bases 1 to 188604)
AUTHORS Xu, W., Hua, A., Eichler, E. and Roe, B.A.
TITLE Papio anubis BAC Clone rp41-273g19
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 188604)
AUTHORS Xu, W., Hua, A., Eichler, E. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (29-MAR-2002) Department Of Chemistry And Biochemistry,
          The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
          OK 73019, USA
REFERENCE 3 (bases 1 to 188604)
AUTHORS Xu, W., Hua, A., Eichler, E. and Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (11-JUN-2003) Department Of Chemistry And Biochemistry,
          The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
          OK 73019, USA
COMMENT On May 15, 2003 this sequence version replaced gi:28173120.
          ----- Genome Center
          Center: Department Of Chemistry
          The University Of Oklahoma
          Center code:UOKNOR
          * NOTE: This is a 'working draft' sequence. It currently
          * consists of 2 contigs. Gaps between the contigs
          * are represented as runs of N. The order of the pieces
          * is believed to be correct as given, however the sizes
          * of the gaps between them are based on estimates that have
          * been provided by the submitter.
          * This sequence will be replaced
          * by the finished sequence as soon as it is available and
          * the accession number will be preserved.
          * 1 6057: contig of 6057 bp in length
          * 6058 6157: gap of unknown length
          * 6158 188604: Contig of 182447 bp in length.
          Location/Qualifiers
            1..188604
              /organism="Papio anubis"
              /mol_type="genomic DNA"
              /db_xref="taxon:9555"
              /clone="rp41-273g19"

FEATURES
source
```

```
/clone_lib="RPCI - 41 Male (Olive) Baboon BAC Library"
BASE COUNT 53384 a 41769 c 40796 g 52555 t 100 others
ORIGIN

Alignment Scores:
Pred. No.: 2.05e-25 Length: 188604
Score: 300.00 Matches: 79
Percent Similarity: 28.67% Conservative: 3
Best Local Similarity: 27.62% Mismatches: 9
Query Match: 58.37% Indels: 195
DB: 2 Gaps: 1

US-10-045-180A-3 (1-94) x AC116558 (1-188604)
QY 4 LeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGluProLeuGln 23
Db 142657 CTCACCTCTCTCTGTCATTTCTCTGGGGCCCTCCAGGCTGCACAGACGCGCTCCAG 142598
QY 24 AlaArgAlaHisGluMetProAlaGlnIlySerProAlaAspAspGlnAspValVal 43
Db 142597 GCAAGAGCTGATGAGATGCCAGCCAGGAGCAGGCTCCAGCAACAGCACAAGATGTGTC 142538
QY 44 IleTyrPheSerGlyAspSerCysSerLeuGlnVal----- 56
Db 142537 ATTACTTTTTCAGAGATGACAGCTCTCTCTTCAGGTTCCAGGTGAAGACACCGCAT 142478
QY 56 ----- 56
Db 142477 GAAGAGCTACAGACTAGAGGGAAGGAGCGGAGACAGGCTCTGGAATTGGGTCTCAGTGGT 142418
QY 56 ----- 56
Db 142417 ACATGTCACTTAGGTGGCTTTTACTTAACATCTCTGTCTCTGGTTGTCTCATGTCTAAAT 142358
QY 56 ----- 56
Db 142357 GGGATAGAGAACCAATAATAATCTAAGAGATTTTCTTCCAAAACATGATTTCCAAGA 142298
QY 56 ----- 56
Db 142297 TATGACTGAAATTCACCTAGGTTTAAAGATATAAGGAGATGCTACCTAGTTCTCTTCGGATC 142238
QY 56 ----- 56
Db 142237 CAGACAAATAAGCTTAACTATATAGAAAATATTTCCCGTGTCTATGAGGAGTTTAG 142178
QY 56 ----- 56
Db 142177 AACCTGGAGAGAGCCTAAGAGTGTGTCAAGTGTGTGTGATGGGGCAGAGGGCAAA 142118
QY 56 ----- 56
Db 142117 AAAGTGAAGCCAAAGAGAATGAGTCTCGAGTCTCTGTGTGACCAGCAGCTGCTCTGTGTATT 142058
QY 56 ----- 56
Db 142057 TATTCCTATTGACTGAGATTGTTTCTGCTACCGGCTGCAATACAGCAACATCACTCATC 141998
QY 56 ----- 56
Db 141997 AGTCAACAGGTGACTTCTCCAGAGATTCCTTACCATCCACTCGGACCCCATACTGAAT 141938
QY 57 -----ProGlySerThrIlySGlyLeuIleCysHisCysArgVa 69
Db 141937 TTCTGATGCTCTCTCTGTTTCCCAAGGCTCAACGAAGGGCTTGAGCTGCCATTGCAGAGT 141878
QY 69 lleuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIleLeuGlyGluArgTy 89
Db 141877 ACTTACTAGCCTTTTGTGAGAACATCTTGGTGGACCTCTTTCATCCATGGTGAAGCTC 141818
QY 89 rProIleCysCysTyr 94
Db 141817 CCCAATCTGCTGCTAC 141802
```



```

QY 21 ProLeuGlnAlaAArgAlaHisGlu---MetProAlaGlnLysGlnProProAlaAspAsp 39
DB 112 CCACCTCCAGGCAAGAACTGACAGAGCTACTGAGCCCGAGGAGCATTCACACAGACAAC 171
QY 40 GlnAspValValIleTyPheSerGlyAspAspSerCysSerLeuGln-----ValPro 57
DB 172 CCAGAAGTGGTGTTCCTTCCTCGCTGGGATGAAGCTTGGCTCCAAAGGATTCAGTCCCA 231
QY 58 GlySerThrLysGlyLeuLeuLeuCysHisCysArgValLeuTyCysIlePheGlyGluHis 77
DB 232 GCCTTAAGGAAACATGCGCTGCTATTGCAGAAATACCAGCGTGTCTTAGCAGGAGAACGT 291
QY 78 LeuGlyGlyThrCysPheIleLeuGlyGluArgTyProIleCysCys 93
DB 292 CGCTATGGAACCTGCTTCTACCTGGGAAGAGTCTGGGCATTCTGCTGC 339

RESULT 10
AF184160 435 bp mRNA linear PRI 14-NOV-1999
LOCUS Macaca mulatta alpha-defensin 1A (MNP1A) mRNA, complete cds.
DEFINITION
ACCESSION AF184160
VERSION AF184160.1 GI:6409289
KEYWORDS
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta

REFERENCE
AUTHORS Zhao, C., Nguyen, T. and Lehrer, R.I.
TITLE cDNA cloning of three alpha-defensins and three demidefensins from rhesus monkey bone marrow
JOURNAL Unpublished
AUTHORS Zhao, C., Nguyen, T. and Lehrer, R.I.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1999) Medicine, UCLA, 10833 Le Conte Avenue, Los Angeles, CA 90095, USA

FEATURES
source
1..435
Location/Qualifiers
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
gene 1..435
CDS 13..303
/gene="MNP1A"
/codon_start=1
/product="alpha-defensin 1A"
/protein_id="AAF07926.1"
/db_xref="GI:6409290"
/translation="MRTLAILAAILLVALQAOEPLQARTDEATRAAQEIPTDNPVV
VSLAWDES LAPKDSVPGLRKNMACYCRIPLACLAGERRYGTCFYLGRVWAFCC"

BASE COUNT 120 a 113 c 104 g 98 t
ORIGIN
Alignment Scores:
Pred. No.: 4..8e-19 Length: 435
Score: 225.50 Matches: 48
Percent Similarity: 65.62% Conservative: 15
Best Local Similarity: 50.00% Mismatches: 30
Query Match: 43.87% Indels: 3
DB: 9 Gaps: 2

US-10-045-180A-3 (1-94) x AF184160 (1-435)

QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
DB 13 ATGAGGACCTCGCCATCTCTGTCGCAATTCCTCGTGGCCCTGAGGCCCGAGGCTGAG 72
QY 21 ProLeuGlnAlaAArgAlaHisGlu---MetProAlaGlnLysGlnProProAlaAspAsp 39
DB 73 CCACCTCCAGGCAAGAACTGACAGAGCTACTGAGCCCGAGGAGCATTCACACAGACAAC 132

QY 21 ProLeuGlnAlaAArgAlaHisGlu---MetProAlaGlnLysGlnProProAlaAspAsp 39
DB 73 CCACCTCCAGGCAAGAACTGACAGAGCTACTGAGCCCGAGGAGCATTCACACAGACAAC 132

```

```

QY 40 GlnAspValValIleTyPheSerGlyAspAspSerCysSerLeuGln-----ValPro 57
DB 133 CCAGAAGTGGTGTTCCTTCCTCGCTGGGATGAAGCTTGGCTCCAAAGGATTCAGTCCCA 192
QY 58 GlySerThrLysGlyLeuLeuLeuCysHisCysArgValLeuTyCysIlePheGlyGluHis 77
DB 193 GCCTTAAGGAAACATGCGCTGCTATTGCAGAAATACCAGCGTGTCTTAGCAGGAGAACGT 252
QY 78 LeuGlyGlyThrCysPheIleLeuGlyGluArgTyProIleCysCys 93
DB 253 CGCTATGGAACCTGCTTCTACCTGGGAAGAGTCTGGGCATTCTGCTGC 300

RESULT 11
AF184159 433 bp mRNA linear PRI 14-NOV-1999
LOCUS Macaca mulatta alpha-defensin 1 (MNP1) mRNA, complete cds.
DEFINITION
ACCESSION AF184159
VERSION AF184159.1 GI:6409287
KEYWORDS
SOURCE Macaca mulatta (rhesus monkey)
ORGANISM Macaca mulatta

REFERENCE
AUTHORS Zhao, C., Nguyen, T. and Lehrer, R.I.
TITLE cDNA cloning of three alpha-defensins and three demidefensins from rhesus monkey bone marrow
JOURNAL Unpublished
AUTHORS Zhao, C., Nguyen, T. and Lehrer, R.I.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-1999) Medicine, UCLA, 10833 Le Conte Avenue, Los Angeles, CA 90095, USA

FEATURES
source
1..433
Location/Qualifiers
/organism="Macaca mulatta"
/mol_type="mRNA"
/db_xref="taxon:9544"
gene 1..433
CDS 13..303
/gene="MNP1"
/codon_start=1
/product="alpha-defensin 1"
/protein_id="AAF07925.1"
/db_xref="GI:6409288"
/translation="MRTLAILAAILLVALQAOEPLQARTDEATAAQEQIPTDNPVV
VSLAWDES LAPKDSVPGLRKNMACYCRIPLACLAGERRYGTCFYMGRVWAFCC"

BASE COUNT 121 a 113 c 102 g 97 t
ORIGIN
Alignment Scores:
Pred. No.: 8..52e-19 Length: 433
Score: 223.50 Matches: 47
Percent Similarity: 65.62% Conservative: 16
Best Local Similarity: 48.96% Mismatches: 30
Query Match: 43.48% Indels: 3
DB: 9 Gaps: 2

US-10-045-180A-3 (1-94) x AF184159 (1-433)

QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
DB 13 ATGAGGACCTCGCCATCTCTGTCGCAATTCCTCGTGGCCCTGAGGCCCGAGGCTGAG 72
QY 21 ProLeuGlnAlaAArgAlaHisGlu---MetProAlaGlnLysGlnProProAlaAspAsp 39
DB 73 CCACCTCCAGGCAAGAACTGACAGAGCTACTGAGCCCGAGGAGCATTCACACAGACAAC 132
QY 40 GlnAspValValIleTyPheSerGlyAspAspSerCysSerLeuGln-----ValPro 57
DB 73 CCACCTCCAGGCAAGAACTGACAGAGCTACTGAGCCCGAGGAGCATTCACACAGACAAC 132

```



```
variation 390 /replace="c"
polyA_signal 424..429
BASE COUNT 124 a 118 c 98 g 112 t
ORIGIN
Alignment Scores: 2.46e-18 Length: 452
Pred. No.: 220.00 Matches: 49
Score: 220.00 Conservative: 14
Percent Similarity: 63.00% Mismatches: 29
Best Local Similarity: 49.00% Indels: 8
Query Match: 42.80% Gaps: 2
DB: 9

US-10-045-180A-3 (1-94) x HUMDFNSIX (1-452)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
DB 19 ATGAGAACCTCACCATCTCACTGCTTCTCTCGTGGCCCTCCAGGCCAAGCTGAG 78
QY 21 -----ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLys 33
DB 79 CCACCTCCAAAGCTGAGGATGATCCACTGCAGGCCAAAGCTTATGAGGCTGATGCCCAGGAG 138
QY 34 GlnProProAlaAspGlnAspValValIleTyrPheSerGlyAspAspSerCysSer 53
DB 139 CAGCGTGGGGCAATGACGAGGACTTTCGCGCTCTCTTTCAGAGGATGCAAGCTCAAGT 198
QY 54 LeuGlnValProGlySerThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIle 73
DB 199 CTTAGAGTTTGGGCTCAACAGGGCTTTCACCTGCCATTGGCAGA---AGTCTCTGTAT 255
QY 74 PheGlyGluHisLeuGlyGlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
DB 256 TCAACAGAAATATTCCTATGGGACCTGCAGCTGCATGGGTATTAAACACAGATTCTGCTGC 315

RESULT 15
HUMDEF1A 448 bp mRNA linear PRI 07-NOV-1994
LOCUS Human neutrophil peptide (defensin) 1 mRNA, complete cds.
DEFINITION M21130 M22160
ACCESSION M21130 M22160
VERSION M21130.1 GI:181526
KEYWORDS cytotoxic protein; defensin 1; microcidal protein; neutrophil peptide.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 448)
Daher, K.A., Lehrer, R.I., Ganz, T. and Kronenberg, M.
Isolation and characterization of human defensin cDNA clones
Proc. Natl. Acad. Sci. U.S.A. 85 (19), 7327-7331 (1988)
MEDLINE 89017191
PUBMED 3174637
COMMENT Original source text: Human leukocyte (myeloid cell line HL-60), cDNA to mRNA, clones HNP1a and HNP1b. Draft entry and computer-readable sequence [1] kindly submitted by K.Daher, 25-NOV-1988. Submission form for [1] kindly provided by M.Kronenberg, 01-JAN-1989.
FEATURES
source Location/Qualifiers
1..448
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/map="8p23-p22"
1..448
/gene="DEF1"
<1..448
/gene="DEF1"
/product="neutrophil peptide 1 mRNA"
51..335
/gene="DEF1"
```

```
/note="neutrophil peptide 1 precursor"
/codon_start=1
/protein_id="AAA52302.1"
/db_xref="GI:181527"
/db_xref="GDB:G00-120-745"
/translation="MRTLAIALLVLAQAEPLOARDEVAAPQIAADIPEVV
VSLAWDSLAPKHPGSKNMACYCRIPACIAGERRYGTCTIYQGRWAFCC"
51..107
/gene="DEF1"
/note="neutrophil peptide 1 signal peptide"
243..332
/gene="DEF1"
/product="neutrophil peptide 1"
BASE COUNT 106 a 127 c 110 g 105 t
ORIGIN Unreported.
Alignment Scores: 3.77e-18 Length: 448
Pred. No.: 218.50 Matches: 47
Score: 64.89% Conservative: 14
Percent Similarity: 50.00% Mismatches: 32
Best Local Similarity: 42.51% Indels: 1
Query Match: 9 Gaps: 1
DB: 9
US-10-045-180A-3 (1-94) x HUMDEF1A (1-448)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
DB 51 ATGAGGACCTCGCCATCTTGTGCCATTCTCTGTGGCCCTGCAGGCCCAGGCTGAG 110
QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProAlaAspAsp 39
DB 111 CCACCTCCAGGCAAGAGCTGATGAGTTCGTGAGCCCGGAGCAGATTGCGAGCGGACATC 170
QY 40 GlnAspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySer 59
DB 171 CCAGAGTGTGTGTTTCCCTTCATGGGACGAAAGCTTGGCTCCAAAGCATCCAGGCTCA 230
QY 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
DB 231 AGGMAAACATGGCGCTGCTATTGCAGAAATACCCAGGCTGCATTGCAGGAGACGTCGCTAT 290
QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
DB 291 GGAACCTGCATCTACCAGGGAAGACTCTGGGCACTTCTGCTGC 332
Search completed: December 17, 2003, 16:36:39
Job time : 2738.5 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 3286.5 Seconds
(without alignments)
695.153 Million cell updates/sec

Title: US-10-045-180A-3
Perfect score: 514
Sequence: 1 MRTLISAFLLVALQWAE.....GEHLGGTCFILGERYPICCY 94

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO spool /US10045180/runat.17122003.145024.12540/app.query.fasta.1.860
-DB=BST -QPWT=fastap -SUPPIX=first -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180@cgn.1.1.6100 @runat.17122003.145024.12540 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hrc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rod:*
26: em_ges_phg:*
27: em_ges_vri:*
28: gb_ges1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	227	44.2	422	13	EX093382	EX093382 BX093382
C 2	227	44.2	634	13	BU616655	BU616655 UI-H-DF0-
C 3	218.5	42.5	349	12	BI021617	BI021617 CM3-MT034
C 4	218.5	42.5	370	10	BF893383	BF893383 QV3-MT012
5	218.5	42.5	372	10	BF893378	BF893378 QV3-MT012
6	218.5	42.5	379	14	CO1361	CO1361 HUMGS000834
7	218.5	42.5	387	12	BI021837	BI021837 RCS-MT025
C 8	218.5	42.5	389	12	BI021710	BI021710 CM3-MT034
C 9	218.5	42.5	396	10	BF893221	BF893221 QV3-MT012
C 10	218.5	42.5	413	9	AW468012	AW468012 he30od11.x
C 11	218.5	42.5	419	10	BF893435	BF893435 QV3-MT012
C 12	218.5	42.5	424	10	BF893232	BF893232 QV3-MT012
C 13	218.5	42.5	425	10	BF893231	BF893231 QV3-MT012
C 14	218.5	42.5	430	9	AI684407	AI684407 wa72b06.x
C 15	218.5	42.5	430	9	AI806882	AI806882 wf36d08.x
C 16	218.5	42.5	433	10	BF893440	BF893440 QV3-MT012
C 17	218.5	42.5	449	10	BF893226	BF893226 QV3-MT012
C 18	218.5	42.5	450	10	BF893220	BF893220 QV3-MT012
C 19	218.5	42.5	465	9	AI312619	AI312619 qp77c06.x
C 20	218.5	42.5	472	10	BF893438	BF893438 QV3-MT012
C 21	218.5	42.5	475	10	BF893227	BF893227 QV3-MT012
C 22	218.5	42.5	486	9	AV741509	AV741509 AV741509
C 23	218.5	42.5	500	13	BU616031	BU616031 UI-H-DF0-
C 24	218.5	42.5	501	13	BU615878	BU615878 UI-H-DF0-
C 25	218.5	42.5	503	12	BM989070	BM989070 UI-H-DF0-
C 26	218.5	42.5	503	13	BQ188934	BQ188934 UI-E-EJ1-
C 27	218.5	42.5	505	13	BQ182900	BQ182900 UI-H-EU0-
C 28	218.5	42.5	506	13	BQ446731	BQ446731 UI-H-EU1-
C 29	218.5	42.5	506	14	CA427527	CA427527 UI-H-DF0-
C 30	218.5	42.5	508	13	BU616082	BU616082 UI-H-DF0-
C 31	218.5	42.5	601	9	AI064980	AI064980 HA0805 Hu
C 32	216.5	42.1	349	12	BI021835	BI021835 RCS-MT025
C 33	216.5	42.1	500	12	BM914478	BM914478 AGENCOURT
C 34	216.5	42.1	504	13	BQ183657	BQ183657 UI-H-EU0-
C 35	216.5	42.1	555	9	AL697992	AL697992 DKF2p686L
C 36	216	42.0	542	9	AI250799	AI250799 qi36g07.x
C 37	214.5	41.7	414	9	AW467613	AW467613 he22d02.x
C 38	214.5	41.7	441	10	BF757978	BF757978 QV3-CT055
C 39	214.5	41.7	461	9	AI091666	AI091666 ool9h12.x
C 40	214.5	41.7	473	9	AW236289	AW236289 xn72d07.x
C 41	214.5	41.7	480	9	AW236257	AW236257 xn71h04.x
C 42	214.5	41.7	506	13	BQ183268	BQ183268 UI-H-EU0-
C 43	213.5	41.5	421	9	AW291977	AW291977 UI-H-BI2-
C 44	213.5	41.5	461	10	BF906299	BF906299 MR3-MT032
C 45	212.5	41.3	411	9	AW467234	AW467234 he07h02.x

ALIGNMENTS

RESULT 1
EX093382
LOCUS
DEFINITION
BX099382 Scores Nhmpu_Sl Homo sapiens cDNA clone IMAGp988H134549 ;
IMAGE:1858620, mRNA sequence.
ACCESSION
BX099382
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 422)

AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M., Radelof, U., Schneider, D. and Korn, B.
TITLE Human Unigeneset - RZPD3
JOURNAL Unpublished
COMMENT RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 360, D-69120 Heidelberg, Germany
RZPD; IMAGp998H134549.
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
Human Unigeneset - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACAGGAACAGCTATGAC.
FEATURES
source
1. .422
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998H134549 ; IMAGE:1858620"
/tissue type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab host="DH10B"
/clone_lib="Soares NhMPU S1"
/notes="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBHM, pregnant uterus NBHPU, and fetal heart NBH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
BASE COUNT 83 a 115 c 115 g 107 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 1.93e-15 Length: 422
Score: 227.00 Matches: 46
Percent Similarity: 64.52% Conservative: 14
Best Local Similarity: 49.46% Mismatches: 33
Query Match: 44.16% Indels: 0
DB: 13 Gaps: 0
US-10-045-180A-3 (1-94) x BX099382 (1-422)
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db 70 ATGAGGATTATCGCCCTCTCGCTGCTATTCTCTTGTTAGCCCTCCAGGTCGCGCAGGC 129
QY 21 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAspGln 40
Db 130 CCACCTCAGGCAGAGGTGATGAGGCTCCAGGCAGGAGCAGCGTGGCCAGAGACCAG 189
QY 41 AspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySerThr 60
Db 190 GACATATCTATTTCTTTGATGGGATAAAGCTCTGCTCTTCAGGTTTCAGGCTCACAA 249
QY 61 LysGlyLeuLeuCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGly 80
Db 250 AGGGGCATGCTGCTCTTCAGATTAGTATTCTCGCGCGAACAAGAACTTCGTGTGGG 309
QY 81 ThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
Db 310 AACTGCCTCATTTGGTGGTGGTTCACATCTACTGCTGC 348

RESULT 2
BU616655/c
LOCUS
DEFINITION UI-H-DF0-ben-n-14-0-UI-s1 NCI CGAP DF0 Homo sapiens cDNA clone
UI-H-DF0-ben-n-14-0-UI 3', mRNA sequence.
ACCESSION BU616655
VERSION 1 GI:23282870
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 634)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.
FEATURES
source
1. .634
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-DF0-ben-n-14-0-UI"
/tissue type="Subchondral Bone"
/dev_stage="Adult"
/lab host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP DF0"
/notes="Organ: Bone; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCI CGAP DF0 is a cDNA library containing the following tissue(s): Subchondral Bone. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GTTAAGCGTC.
TAG LIB=UI-H-DF0
TAG_TISSUE=subchondral bone
TAG_SEQ=GTTAAGCGTC"
BASE COUNT 170 a 151 c 158 g 155 t
ORIGIN
Alignment Scores:
Pred. No.: 3.38e-15 Length: 634
Score: 227.00 Matches: 46
Percent Similarity: 64.52% Conservative: 14
Best Local Similarity: 49.46% Mismatches: 33
Query Match: 44.16% Indels: 0
DB: 13 Gaps: 0
US-10-045-180A-3 (1-94) x BU616655 (1-634)
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db 509 ATGAGGATTATCGCCCTCTCGCTGCTATTCTCTTGTTAGCCCTCCAGGTCGCGCAGGC 450
QY 21 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAspGln 40
Db 948 AACTGCCTCATTTGGTGGTGGTTCACATCTACTGCTGC 980

sequences of the clones represented by this GS sequences is also found there.

FEATURES	source	Location/Qualifiers
	1. .379	
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/dev_stage="adult"
		/clone_lib="Human adult (K.Okubo)"
		/notes="One or more human adult tissue"
BASE COUNT	90 a 102 c 100 g	82 t 5 others
ORIGIN		

Alignment Scores:		
Pred. No.:	1,478-14	Length:
Score:	218.50	Matches:
Percent Similarity:	64.89%	Conservative:
Best Local Similarity:	50.00%	Mismatches:
Query Match:	42.51%	Indels:
DB:	14	Gaps:
		14
		379

US-10-045-180A-3 (1-94) x C01361 (1-379)

QY	1	MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu	20
Db	18	ATGAGGACCTCGCATCTGTTGCCATTCTCTGTGGCCCTGCAGGCCAGGCTGAG	77
QY	21	ProLeuGlnAlaArgAlaHisGluMet--ProAlaGlnIlysGlnProProAlaAspAsp	39
Db	78	CCACTCCAGGCAAGAGCTGATGAGGTTCTCTGCAGCCCGGAGCGAGATTGCAGCGGCAATC	137
QY	40	GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlyIys	59
Db	138	CCAGAAGTCGTGTGTTCTCCTTCATGCGCAGCAAAAGCTTGCTCCAAAGCATCCAGGCTCA	197
QY	60	ThrIlysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGlnHisLeuGly	79
Db	198	AGGAAACATCGCTGCTATTATGCAATACACAGGTGCATTGCAGAGAAACGCTCGCTAT	257
QY	80	GlyThrCysPheIleLeuGlyGlyIuArgTyrProIleCysCys	93
Db	258	GGNACTGTCATCTACGAGGAAGACTCTGGCATTCTGCTGC	299

RESULT 7	BI021837	LOCUS	BI021837	387 bp	mRNA	linear	EST 14-JUN-2001
DEFINITION	RC5-WT0259-170101-011-B02 MT0259 Homo sapiens cDNA, mRNA sequence.						
ACCESSION	BI021837						
VERSION	BI021837.1 GI:14428467						

KEYWORDS	SOURCE	EST.	ORGANISM
		Homo sapiens	(human)
		Homo sapiens	

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 387)
 DNAS Netto,B., Garcia Correa,R., Vertjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 samples, A.U.
 sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800

20351000
Contact: Simpson A. J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SF,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
This sequence was derived from the
Project. This entry can be seen in
<http://www.ludwig.org.br/scripts/170101-011-B02&t3=2001-01-17&t4=1>
Seq primer: puc 18 forward
High quality sequence stop: 387.
Location/Qualifiers
1. .387
FEATURES
source

FEATURES
source

BASE COUNT	90 a	102 c	100 g	82 t	5 others
ORIGIN					

Alignment Scores:		
Pred. No.:	1,478-14	Length:
Score:	218.50	Matches:
Percent Similarity:	64.89%	Conservative:
Best Local Similarity:	50.00%	Mismatches:
Query Match:	42.51%	Indels:
DB:	14	Gaps:
		14
		379

US-10-045-180A-3 (1-94) x C01361 (1-379)

QY	1	MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu	20
Db	18	ATGAGGACCTCGCATCTGTTGCCATTCTCTGTGGCCCTGCAGGCCAGGCTGAG	77
QY	21	ProLeuGlnAlaArgAlaHisGluMet--ProAlaGlnIlysGlnProProAlaAspAsp	39
Db	78	CCACTCCAGGCAAGAGCTGATGAGGTTCTCGACGCCGCGAGCGAGATTGCGCGGCAATC	137
QY	40	GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlyIys	59
Db	138	CCAGAAGTCGTGTGTTCTCCTTCATCGGCGCAAAAGCTTGCGTCCAAAGCATCCAGGCTCA	197
QY	60	ThrIlysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGlnHisLeuGly	79
Db	198	AGGAAACATCGCTGCTATTGTCAGAAATACACGCTGCATTGCAGAGAAACGCTCGCTAT	257
QY	80	GlyThrCysPheIleLeuGlyGlyIuArgTyrProIleCysCys	93
Db	258	GGNACTGTCATCTACGAGGAAGACTCTGGCATTCTGCTGC	299

RESULT 7	BI021837	LOCUS	BI021837	387 bp	mRNA	linear	EST 14-JUN-2001
DEFINITION	RC5-WT0259-170101-011-B02 MT0259 Homo sapiens cDNA, mRNA sequence.						
ACCESSION	BI021837						
VERSION	BI021837.1 GI:14428467						

KEYWORDS	SOURCE	EST.	ORGANISM
		Homo sapiens	(human)
		Homo sapiens	

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 387)
 DNAS Netto,B., Garcia Correa,R., Vertjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,
 M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
 samples, A.U.
 sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800

2035000
Contact: Simpson A. J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001

```

Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=RC5&t2=RC5-MT0259-
170301-011-B02&t3=2001-01-17&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 387.
Location/Qualifiers
1..387
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0259"
/note="Organ: narrow; Vector: puc18; Site_1: SmaI; Site_2:
SmaI; A mini-library was made by cloning products derived
from ORESTS PCR (U.S. Letters Patent application No. 196
,716 - Ludwig Institute for Cancer Research) profiles
into the puc 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."

```

BASE COUNT	96 a	104 c	104 g	83 t
ORIGIN	stringency conditions."			

Alignment Scores:		
Pred. No.:	1.51e-14	387
Score:	218.50	47
Percent Similarity:	64.89%	Conservative: 14
Best Local Similarity:	50.00%	Mismatches: 32
Query Match:	42.51%	Indels: 1
DB:	12	Gaps: 1

US-10-045-180A-3 (1-94) x BI021837 (1-387)

Qy 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
20 ATGAGGACCTCGCCATCCTTGTGTGCCATCTCTGTGTGGCCCTCGAGGGCCAGGCTGAG 79
Db

[illegible]

Qy 40 GlnAspValValIleTy-PheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
 Db 140 CCAGAAAGTGGTTTTCCTTTCATGGGACGAAAGCTTGGTCCCAAGCATCCAGGCTCA 199

Qy 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
 Db 200 AGGAAACATGGCGCTGCTATTGCAGAAATACCCAGCGTGCATTGCAGGAGAACGTGCGTAT 259

80 GlyThrCysPheIleLeuGlyArgTyrProIleCysCys 93
|||
260 GGAACCTGCATCTACCAGGGAAGACTCTGGGCAATTTCGTGC 301

RESULT 8
BI021710/c

LOCUS	BI021710	389 bp	mRNA	linear	EST 14-JUN-2001
DEFINITION	CH3-MT0343-170101-666-f02 MT0343		Homo sapiens	CDNA, mRNA sequence.	

ACCESSION BI021710
VERSION BI021710.1 GI:14428340
KEYWORDS EST

KEYWORDS	SOURCE	EST.	Homo sapiens (human)
ORGANISM			Homo sapiens

ORGANISM *Homo sapiens*

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE 1 (bases 1 to 389)

I. LUDAS, E. T. 10-3591
 REFERENCES
 DÍAS NETO, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
 M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.
 TITLE
 Shotgun sequencing of the human transcriptome with ORF expressed

TITL8

sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=CM3&t2=CM3-MT0343-
170101-666-f02&t3=2001-01-17&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 52
High quality sequence stop: 389.
Location/Qualifiers
1. .389
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0343"
/note="Organ: marrow; Vector: puc18; Site 1: Smal; Site 2:
Smal; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 81 a 106 c 106 g 96 t
ORIGIN
Alignment Scores:
Pred. No.: 1.52e-14 Length: 389
Score: 218.50 Matches: 47
Percent Similarity: 64.89% Conservative: 14
Best Local Similarity: 50.00% Mismatches: 32
Query Match: 42.51% Indels: 1
DB: 12 Gaps: 1
US-10-045-180A-3 (1-94) x BI021710 (1-389)
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db 351 ATGAGGACCTCGCCATCTCTGTCATTCCTCTGTCGCTCCCTGAGGCCCGAGCTGAG 292
QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAsp 39
Db 291 CCAGTCCAGGACAGAGCTGATGAGTTGCTGCAGCCCGGAGCAGATGTCAGCGGACATC 232
QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
Db 231 CCAGAGTGGTGTGTTCCCTTCATGTCGACGACGAAAGCTTGCTCCAAAGCATCCAGGCTCA 172
QY 60 ThrIysGlyLeuLeuCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
Db 171 AGGAAAACATGGCCTGCTCTATTCAGAAATCCAGACGCTGATTCAGGAGAACGTCGCTAT 112
QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
Db 111 GGAACCTGCATCTACACGAGGAGAGACTCTGGGCATCTTCGCTGC 70
RESULT 9
BF893221/c
LOCUS BF893221 396 bp mRNA linear EST 18-JAN-2001
DEFINITION QV3-WT0129-111100-419-d01 MT0129 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF893221
VERSION BF893221.1 GI:12284680
KEYWORDS EST.

SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 396)
AUTHORS Dias Neto,E., Garcia Correa,R., Vertovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldmann,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV3&t2=QV3-MT0129-111100-419-d01&t3=2000-11-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 38
High quality sequence stop: 396.
Location/Qualifiers
1. .396
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0129"
/note="Organ: marrow; Vector: puc18; Site 1: Smal; Site 2:
Smal; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 88 a 99 c 111 g 98 t
ORIGIN
Alignment Scores:
Pred. No.: 1.56e-14 Length: 396
Score: 218.50 Matches: 47
Percent Similarity: 64.89% Conservative: 14
Best Local Similarity: 50.00% Mismatches: 32
Query Match: 42.51% Indels: 1
DB: 10 Gaps: 1
US-10-045-180A-3 (1-94) x BF893221 (1-396)
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db 394 ATGAGGACCTCGCCATCTCTGTCATTCCTCTGTCGCTCCCTGAGGCCCGAGCTGAG 335
QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAsp 39
Db 334 CCAGTCCAGGACAGAGCTGATGAGTTGCTGCAGCCCGGAGCAGATTCGACGCGACATC 275
QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
Db 274 CCAGAGTGGTGTGTTCCCTTCATGTCGACGACGAAAGCTTGCTCCAAAGCATCCAGGCTCA 215
QY 60 ThrIysGlyLeuLeuCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
Db 214 AGGAAAACATGGCCTGCTCTATTCAGAAATCCAGACGCTGATTCAGGAGAACGTCGCTAT 155

```

Qy      80 GlyThrCysPheLeuLeuGlyGluArgTyrProIleCysCys 93
Db      154 GGAACCTGCATCTACCGAGGAGACTCTGGGCATTCTGCTGC 113

RESULT 10
AW468012/c
LOCUS      AW468012
DEFINITION he30d11.x1 NCI CGAP CML1 Homo sapiens cDNA clone IMAGE:2920533 3'
            similar to gb:M26602 NEUTROPHIL DEFENSINS 1, 2 AND 3 PRECURSOR
            (HUMAN);, mRNA sequence.
ACCESSION AW468012
VERSION    AW468012.1 GI:7038118
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 413)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Elisabeth Paietta, Jonathan D. Licht, M.D.,
            Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life
            Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The
            I.M.A.G.E. Consortium DNA Sequencing by: Washington University
            Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Seq primer: -40UP from Gibco.
FEATURES   Location/Qualifiers
            source          1..413
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /clone="IMAGE:2920533"
                        /tissue type="myeloid cells, 18 pooled CML cases, BCR/ABL
                        rearrangement positive, includes both chronic phase and
                        myeloid blast crisis"
                        /lab host="DH10B"
                        /clone lib="NCI CGAP CML1"
                        /note="Organ: whole blood; Vector: pCMV-SPORT6; Site 1:
                        SalI; Site 2: NotI; Cloned unidirectionally. Primer:
                        Oligo dT. Library constructed by Life Technologies."
BASE COUNT 95 a 101 c 113 g 103 t
ORIGIN
Alignment Scores:
Pred. No.: 1..65e-14 Length: 413
Score: 218.50 Matches: 47
Percent Similarity: 64.89% Conservativeness: 14
Best Local Similarity: 50.00% Mismatches: 32
Query Match: 42.51% Indels: 1
DB: 9 Gaps: 1

US-10-045-180A-3 (1-94) x AW468012 (1-413)

Qy      1 MetArgThrLeuThrLeuLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db      405 ATGAGGACCTCGCCATCTCTGCTGCCATTCTCTGCTGGCCCTGCAGGCAGGCTGAG 346

Qy      21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnIysGlnProProAlaAsp 39
Db      345 CCATCCAGGCAAGAGCTGTAGTGGTGTCTCAGCCCGGAGCAGATTGCAGCGCATC 286

Qy      40 GlnAspValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySer 59
Db      285 CCAGAGTGTGTGTTTCCTTGGCTGATGGAGCAAGAGCTGGCTCCAAAGCATCCAGGCTCA 226

```

```

Qy      60 ThrIysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
Db      225 AGGAAAACATGGCGCTGCTATTGAGAAATACACGCGTGCATTGCAGGAGACGTCGCTAT 166

Qy      80 GlyThrCysPheLeuLeuGlyGluArgTyrProIleCysCys 93
Db      165 GGAACCTGCATCTACCGAGGAGACTCTGGGCATTCTGCTGC 124

RESULT 11
BF893435/c
LOCUS      BF893435
DEFINITION QV3-MT0129-111100-426-c04 MT0129 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF893435
VERSION    BF893435.1 GI:12284894
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 419)
AUTHORS    Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
            Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
            Goldman, G.H., Carvalho, A.F., Mateukuma, A., Baia, G.S., Simpson, D.H.,
            Brunstein, A., de Oliveira, P.S., Sucher, P., Jongeneel, C.V., O'Hare
            , M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.
TITLE      Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
JOURNAL    Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE    20202663
PubMed     10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-MT0129-
            111100-426-c04&t3=2000-11-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence stop: 419.
FEATURES   Location/Qualifiers
            source          1..419
                        /organism="Homo sapiens"
                        /mol_type="mRNA"
                        /db_xref="taxon:9606"
                        /dev stage="Adult"
                        /clone lib="MT0129"
                        /note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:
                        SmaI; A mini-library was made by cloning products derived
                        from ORESTES PCR (U.S. Letters Patent application No. 196
                        ,716 - Ludwig Institute for Cancer Research) profiles
                        into the pUC 18 vector. Reverse transcription of tissue
                        mRNA and cDNA amplification were performed under low
                        stringency conditions."
BASE COUNT 101 a 103 c 114 g 101 t
ORIGIN
Alignment Scores:
Pred. No.: 1..69e-14 Length: 419
Score: 218.50 Matches: 47
Percent Similarity: 64.89% Conservativeness: 14
Best Local Similarity: 50.00% Mismatches: 32
Query Match: 42.51% Indels: 1
DB: 10 Gaps: 1

US-10-045-180A-3 (1-94) x BF893435 (1-419)

Qy      1 MetArgThrLeuThrLeuLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20

```

```

Db      413 ATGAGACCTCGCCATCTTGTGCTGCTGCTGAGGCCCGAGGTGAG 354
QY      21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAspAsp 39
Db      353 CCACCTCCAGGCAAGAGCTGATGAGTTGCTGCAGCCCGGAGCAGATTGCGAGCGGACATC 294
QY      40 GlnAspValValLleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
Db      293 CCAGAAAGTGTGTTGTTCCCTTGCATGGGAGAAAGCTTGGCTCCAAAGCATCCAGGCTCA 234
QY      60 ThrLysGlyLeuLeuLeuGlnCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
Db      233 AGGAAAACATGCGCTGCTATTGCAGATACCGAGCGTGCAATGCAGGAGAACGTCGCTAT 174
QY      80 GlyThrCysPheLeuLeuGlyGluArgTyrProIleCysCys 93
Db      173 GGAACCTGCATCTACCAAGGAAAGACTCTGGGCATCTTCGTGTC 132

RESULT 12
BF893232/c
LOCUS      424 bp mRNA linear EST 18-JAN-2001
DEFINITION QV3-MT0129-111100-419-h09 MT0129 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF893232
VERSION     BF893232.1 GI:12284691
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 424)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            20202663
            10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-MT0129-
            111100-419-h09&t3=2000-11-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 21
            High quality sequence stop: 424.
            Location/Qualifiers
                1..424
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /dev_stage="Adult"
                /clone_lib="MT0129"
                /note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:
                SmaI; A mini-library was made by cloning products derived
                from ORESTES PCR (U.S. Letters Patent application No. 196
                ,716 - Ludwig Institute for Cancer Research) profiles
                into the puc 18 vector. Reverse transcription of tissue
                mRNA and cDNA amplification were performed under low
                stringency conditions."
                96 a 107 c 119 g 102 t

BASE COUNT
ORIGIN

```

```

Alignment Scores:      1.72e-14      Length:      424
Pred. No.:      218.50      Matches:      47
Score:      64.89%      Conservative:      14
Percent Similarity:      50.00%      Mismatches:      32
Best Local Similarity:      42.51%      Indels:      1
Query Match:      10      Gaps:      1
DB:

US-10-045-180A-3 (1-94) x BF893232 (1-424)

QY      1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db      383 ATGAGACCTCGCCATCTTGTGCTGCTGCTGAGGCCCGAGGTGAG 324
QY      21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAspAsp 39
Db      323 CCACCTCCAGGCAAGAGCTGATGAGTTGCTGCAGCCCGGAGCAGATTGCGAGCGGACATC 264
QY      40 GlnAspValValLleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
Db      263 CCAGAAAGTGTGTTGTTCCCTTGCATGGGAGAAAGCTTGGCTCCAAAGCATCCAGGCTCA 204
QY      60 ThrLysGlyLeuLeuLeuGlnCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
Db      203 AGGAAAACATGCGCTGCTATTGCAGATACCAAGCGTGCAATGCAGGAGAACGTCGCTAT 144
QY      80 GlyThrCysPheLeuLeuGlyGluArgTyrProIleCysCys 93
Db      143 GGAACCTGCATCTACCAAGGAAAGACTCTGGGCATCTTCGTGTC 102

RESULT 13
BF893231/c
LOCUS      425 bp mRNA linear EST 18-JAN-2001
DEFINITION QV3-MT0129-111100-419-h03 MT0129 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BF893231
VERSION     BF893231.1 GI:12284590
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 425)
AUTHORS     Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            20202663
            10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=QV3&t2=QV3-MT0129-
            111100-419-h03&t3=2000-11-11&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 23
            High quality sequence stop: 425.
            Location/Qualifiers
                1..425
                /organism="Homo sapiens"
                /mol_type="mRNA"

FEATURES
            source

```



```
/db xref="taxon:9606"
/dev stage="Adult"
/clone lib="MT0129"
/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2:
SmaI; A mini-library was made by cloning products derived
from ORESTES PCR (U.S. Letters Patent application No. 196
716 - Ludwig Institute for Cancer Research) profiles
into the pUC 18 vector. Reverse transcription of tissue
mRNA and cDNA amplification were performed under low
stringency conditions."
BASE COUNT 96 a 108 c 118 g 103 t
ORIGIN

Alignment Scores:
Pred. No.: 1 72e-14 Length: 425
Score: 218.50 Matches: 47
Percent Similarity: 64.89% Conservative: 14
Best Local Similarity: 50.00% Mismatches: 32
Query Match: 42.51% Indels: 1
DB: 10 Gaps: 1

US-10-045-180A-3 (1-94) x BF893231 (1-425)

Qy 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db 387 ATGAGGACCCCTCGCATCTTCTGTCATTTCTCTGTTGGCCCTGCAGGCCAGGCTGAG 328
Qy 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAsp 39
Db 327 CCATCCAGGACGAGCTGATGAGGTTCTCTGAGCCCGGAGCAGATTGCAGCGACATC 268
Qy 40 GlnAspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySer 59
Db 267 CCAGAGTGGTGTTCCTTCGTCAGGACGAAAGCTTGGCTCCAAAGCATCCAGGCTCA 208
Qy 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
Db 207 AGGAAAACATGGCTGCTATTGCAGATATACAGCGTGATTCGAGGACGTCGCTAT 148
Qy 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
Db 147 GGAACCTGCATCTACAGGGAAGACTCTGGGCATTTCTGCTGC 106

RESULT 14
Al806882/c
LOCUS
DEFINITION
IMAGE:2301683 3', similar to gb:M26602 NEUTROPHIL DEFENSINS 1, 2 AND
3 PRECURSOR (HUMAN); mRNA sequence.
Al806882
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 430)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 526 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .430
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2301683"

FEATURES
source
```

```
/lab host="DH10B"
/clone lib="Soares_NFL_T_GBC_S1"
/note="Organ: pooled; Vector: p773D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBH19W, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo.
BASE COUNT 100 a 106 c 115 g 109 t
ORIGIN

Alignment Scores:
Pred. No.: 1 75e-14 Length: 430
Score: 218.50 Matches: 47
Percent Similarity: 64.89% Conservative: 14
Best Local Similarity: 50.00% Mismatches: 32
Query Match: 42.51% Indels: 1
DB: 9 Gaps: 1

US-10-045-180A-3 (1-94) x AT684407 (1-430)

Qy 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db 404 ATGAGGACCCCTCGCATCTTCTGTCATTTCTCTGTTGGCCCTGCAGGCCAGGCTGAG 345
Qy 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAsp 39
Db 344 CCATCCAGGACGAGCTGATGAGGTTCTCTGAGCCCGGAGCAGATTGCAGCGACATC 285
Qy 40 GlnAspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySer 59
Db 284 CCAGAGTGGTGTTCCTTCGTCAGGACGAAAGCTTGGCTCCAAAGCATCCAGGCTCA 225
Qy 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
Db 224 AGGAAAACATGGCTGCTATTGCAGATATACAGCGTGATTCGAGGACGTCGCTAT 165
Qy 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
Db 164 GGAACCTGCATCTACAGGGAAGACTCTGGGCATTTCTGCTGC 123

RESULT 15
Al806882/c
LOCUS
DEFINITION
IMAGE:2357679 3', similar to gb:M26602 NEUTROPHIL DEFENSINS 1, 2 AND
3 PRECURSOR (HUMAN); mRNA sequence.
Al806882
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 430)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 286 Std Error: 0.00
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .430
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2301683"

FEATURES
source
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 352.5 Seconds
(without alignments)
719.950 Million cell updates/sec

Title: US-10-045-180A-3

Perfect score: 514

Sequence: 1 MRTLTLISAFLLVALQWAE.....GEHLGGTCFILGERYPICCY 94

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q/cygn2_1/USPTO_spool_p/US10045180/runat_17122003_145023_12520/app_query.fasta_1.860
-DB=N Geneseq 19Jun03 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=pto -NORF=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180 -CGN 1 835 -runat_17122003_145023_12520 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	514	100.0	453	20	AA26697
2	514	100.0	453	20	AA26698
3	393.5	76.6	4415	20	AA26696
4	227	44.2	860	25	ABX63473
5	220	42.8	452	14	AAQ53217
6	220	42.8	452	25	ACAS6600
7	218.5	42.5	478	21	AAQ04468
8	218.5	42.5	498	12	AAQ14908
9	218.5	42.5	498	24	ABQ80601
10	218.5	42.5	514	24	ABN59722
11	218.5	42.5	652	24	ABZ11202
12	218.5	42.5	664	24	ABZ11183
13	216.5	42.1	464	21	AAZ20911
14	216.5	42.1	464	21	AAZ34789
15	216.5	42.1	644	21	AAF21446
16	216.5	42.1	209273	21	AAF21437
17	209.5	40.8	299	16	AAZ26106
18	207	40.3	424	14	AAQ53216
19	207	40.3	449	24	ABQ60783
20	200	38.9	411	24	ABQ60009
21	198	38.5	586	24	ABQ60212
22	197.5	38.4	543	25	ABZ36767
23	196	38.1	514	22	AAH57427
24	175.5	34.1	496	25	AAQ51767
25	173.5	33.8	420	17	AAZ30745
26	171.5	33.4	401	17	AAZ30753
27	168.5	32.8	388	17	AAZ30751
28	166.5	32.4	377	17	AAZ30747
29	165.5	32.2	422	17	AAZ30740
30	165.5	32.2	500	22	AAQ66924
31	165.5	32.2	500	25	AAQ51768
32	164.5	32.0	422	17	AAZ30741
33	163.5	31.8	294	22	AAH57220
34	163.5	31.8	422	17	AAZ30739
35	160	31.1	365	17	AAZ30742
36	160	31.1	412	24	ABK39009
37	160	31.1	412	25	ACAL1338
38	160	31.1	412	25	ACA02524
39	159.5	31.0	391	17	AAZ30754
40	158.5	30.8	375	17	AAZ30748
41	157.5	30.6	419	17	AAZ30729
42	157.5	30.6	422	17	AAZ30744
43	157.5	30.6	495	22	AAQ66925
44	157.5	30.6	495	25	AAQ51769
45	154.5	30.1	421	17	AAZ30743

ALIGNMENTS

RESULT 1

AA26697

ID AA26697 standard; cDNA; 453 BP.

AC AA26697;

XX 18-JUN-1999 (first entry)

DT cDNA sequence encoding human defensin (Def-X) protein.

DE Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
KW cytosolic; anticancer; inflammation; tissue repair;
KW endocrine regulation; corticostatic regulation; cancer; melanoma;
KW AIDS; immune deficiency; psoriasis; ss.

XX Homo sapiens.

OS

XX

AC ABX63473;
 XX
 DT 26-FEB-2003 (first entry)
 XX
 DE Human cDNA #473 differentially expressed in activated vascular tissue.
 XX
 KW Human; gene; ss: vascular tissue; cytostatic; atherosclerosis;
 KW cardiant; hypotensive; antidiabetic; gynaecological; vasotropic;
 KW cerebroprotective; gene therapy; vascular disease; cancer; coronary;
 KW artery disease; hypertension; diabetes; pre-eclampsia; restenosis;
 KW ischaemia-reperfusion injury; stroke;
 XX
 OS Homo sapiens.
 XX
 XX US2002137081-A1.
 XX
 XX 26-SEP-2002.
 XX
 XX 08-JAN-2002; 2002US-0044090.
 XX
 XX 28-JUL-2000; 2000US-222469P.
 PR 08-JAN-2001; 2001US-260483P.
 XX
 XX (BAND/) BANDMAN O.
 XX
 XX Bandman O;
 XX
 DR WPI; 2003-110597/10.
 XX
 XX
 PT Combination for diagnosing, staging, treating, or monitoring the
 PT progression of treatment of a vascular disease, e.g. atherosclerosis,
 PT comprises several cDNAs that are differentially expressed in activated
 PT vascular tissue -
 XX
 XX Claim 1; Page -; 18pp; English.
 PS
 SS This invention relates to a combination comprising several cDNAs that
 CC are differentially expressed in activated vascular tissue. The invention
 CC also discloses a high throughput method for detecting differentially
 CC expressed cDNAs in a sample. The cDNAs of the invention may have
 CC antiarteriosclerotic; cytostatic; cardiant; hypotensive; antidiabetic;
 CC gynaecological; vasotropic and cerebroprotective activities and may be
 CC used in gene therapy. The cDNAs of the invention may be used in a
 CC high-throughput methods for detecting differential expression of one or
 CC more cDNAs in a sample, or screening several molecules or compounds to
 CC identify a molecule or compound that specifically binds a cDNA of the
 CC invention. A protein encoded by the cDNA may be used to screen several
 CC molecules or compounds to identify a ligand that specifically binds to
 CC the protein, or to produce or purify an antibody to the protein that can
 CC be used to detect a protein in a sample or purify a natural or
 CC recombinant protein from a sample. The nucleotides may be useful for
 CC diagnosing, staging, treating, or monitoring the progression of
 CC treatment of a vascular disease, e.g. atherosclerosis, cancer, coronary
 CC artery disease, hypertension, diabetes, pre-eclampsia, ischaemia-
 CC reperfusion injury, restenosis, or stroke. The cDNAs can also be used
 CC for large-scale genetic or gene expression analysis of several new
 CC nucleic acid molecules. Antibodies to the proteins encoded by the
 CC cDNAs are useful for diagnosing pre-pathologic disorders, and chronic
 CC or acute diseases associated with abnormalities in the expression,
 CC amount or distribution of the protein. The present sequence
 CC represents a cDNA of the invention that is differentially expressed in
 CC activated vascular tissue.
 CC Note: The sequence data for this patent did not form part of the
 CC specification, but was obtained in electronic format directly from USPTO
 CC at <http://seqdata.uspto.gov/sequence.html?DocID=20020137081>.
 XX
 SQ Sequence 860 BP; 216 A; 211 C; 192 G; 241 T; 0 other;

Alignment Scores:

Pred. No.:	2,87e-18	Length:	860
Score:	227.00	Matches:	46
Percent Similarity:	64.52%	Conservative:	14
Best Local Similarity:	49.46%	Mismatches:	33

Query Match:	44.16%	Indels:	0
DB:	25	Gaps:	0

US-10-045-180A-3 (1-94) x ABX63473 (1-860)

QY	1	MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu	20
DB	76	ATGAGGATTATCGCCCTCCCTGCTGCTATTTCTTGTAGCCCTCCAGGTCGGGAGGC	135
QY	21	ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaLeuAspGln	40
DB	136	CACTCCAGCAGAGGTGATGAGGCTCCAGGCGAGCGTGGGCGAAGACCAAGACCAAG	195
QY	41	AspValValIleTyrPheSerGlyAspSerGlySerLeuGlnValProGlySerThr	60
DB	196	GACATATCTATTCTTCATGGGATATAAGCTCTGCTCTTCAGGTTTCAGGCTCAACA	255
QY	61	LysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGly	80
DB	256	AGGGCATGCTGCTCTTCAGATTTCTGCCGCGAAGCACTCTGCTGTGGG	315
QY	81	ThrCysPheIleLeuGlyGluArgTyrProIleCysCys	93
DB	316	AACTGCTCATTTGGTGTGAGTTTCACATCTGCTGC	354

RESULT 5
 AAQ53217
 ID AAQ53217 standard; cDNA; 452 BP.
 XX
 AC AAQ53217;
 XX
 DT 25-MAR-2003 (updated)
 DT 17-JUN-1994 (first entry)
 XX
 DE Sequence encoding gastrointestinal defensin (GID) peptide called
 DE human defensin 6.
 XX
 KW Gastrointestinal defensin peptide; GID; pharmaceutical; Paneth
 KW cell; antimicrobial; anti-inflammatory; diagnosis;
 KW contact disinfectant; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 19..321
 FT /*tag= a
 XX
 PN WO9324513-A1.
 XX
 PD 09-DEC-1993.
 XX
 PF 18-MAY-1993; 93WO-US04740.
 XX
 PR 22-MAY-1992; 92US-0889232.
 XX
 PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
 XX
 PI Bevins CL, Jones DE;
 XX
 DR WPI; 1993-405719/50.
 DR P-PSDB; AAR44819.
 XX
 PT Gastrointestinal defensin peptide(s) - useful as antimicrobial
 PT and anti-inflammatory agents and for detecting gastrointestinal
 PT disorders
 XX
 PS Claim 4; page 48-49; 97pp; English.
 XX
 CC A probe (D5' oligo) having the sequence given in AAQ53219 (claimed)
 CC was used to screen a population of clones to identify novel defensin
 CC defensin peptides. Pref. the libraries were human genomic and cDNA
 CC libraries. Hybridisation and partial sequence analysis of the
 CC identified clones contained previously characterised myeloid derived

QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
 DB 214 CCAGAGTGGTGTTCCTTCATGGGACGAAGCTTGGCTCCAAAGCATCCAGGCTCA 273
 QY 60 ThrLysGlyLeuLeuCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
 DB 274 AGGAAACATGGCTGCTATTGCAGAAATACAGCGTGCATTGCAGGAGACGTCGCTAT 333
 QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
 DB 334 GGAACCTGCATCTACAGGAAGACTCTGGCAATCTGCTGC 375
 RESULT 9
 ID ABQ80601 standard; DNA; 498 BP.
 AC ABQ80601;
 DT 11-NOV-2002 (first entry)
 DE Human defensin 1 coding sequence.
 KW Antiallergic; allergic disease; allergy; defensin 1; atopic dermatitis;
 KW human; gene; ds.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 94..378
 FT /*tag= a
 FT /product= "Human defensin 1"
 PN WO200264832-A1.
 XX 22-AUG-2002.
 XX 13-FEB-2002; 2002WO-JP01193.
 XX 14-FEB-2001; 2001JP-0036362.
 XX (GENO-) GENOX RES INC.
 XX (NIGE-) JAPAN GEN AGENCY NATION.
 XX Sugita Y, Heishi M, Kagaya S, Gunji S, Saito H;
 WPI; 2002-674884/72.
 P-PSDB; ABB98494.
 XX Examining allergic diseases by changes in expression levels of
 PT defensin1 gene in peripheral blood monocytes as indication, also
 PT applicable in screening compounds for treating of allergic diseases
 PT e.g. atopic dermatitis -
 PS Disclosure; Page 45-46; 57pp; Japanese.
 XX The present invention relates to a method for examining allergic
 CC diseases. The method comprises: (a) determining the expression level of
 CC defensin 1 gene in the biological sample from a patient; and (b)
 CC comparing the expression level with that in the sample of a healthy
 CC individual. The method is for examining allergic diseases particularly
 CC atopic dermatitis and its diagnosis, which is also applicable in
 CC screening candidate compounds for remedies. The present sequence is the
 CC coding sequence for human defensin 1.
 SQ Sequence 498 BP; 118 A; 134 C; 122 G; 111 T; 13 other;
 Alignment Scores:
 Pred. No.: 1.55e-17 Length: 498
 Score: 218.50 Matches: 47
 Percent Similarity: 64.89% Conservative: 14
 Best Local Similarity: 50.00% Mismatches: 32
 Query Match: 42.51% Indels: 1
 DB: 24 Gaps: 1

US-10-045-180A-3 (1-94) x ABQ80601 (1-498)
 QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
 DB 94 ATGAGGACCTCGCCATCTTCTGCTGCCATCTCTCTGGTGGCCCTGCAGGCCGAGCTGAG 153
 QY 21 ProLeuGlnAlaArgAlaHisGluWet---ProAlaGlnLysGlnProProAlaAspAsp 39
 DB 154 CCACTCCAGGCAAGAGCTGATGAGGTTGCTGCAGCCCGGAGCAGATTGCAGCGGACATC 213
 QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
 DB 214 CCAGAGTGGTGTTCCTTCATGGGACGAAGCTTGGCTCCAAAGCATCCAGGCTCA 273
 QY 60 ThrLysGlyLeuLeuCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
 DB 274 AGGAAACATGGCTGCTATTGCAGAAATACAGCGTGCATTGCAGGAGACGTCGCTAT 333
 QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
 DB 334 GGAACCTGCATCTACAGGAAGACTCTGGCAATCTGCTGC 375
 RESULT 10
 ABN59722
 ID ABN59722 standard; cDNA; 514 BP.
 AC ABN59722;
 DT 28-JUN-2002 (first entry)
 DE Novel human coding sequence SEQ ID NO: 133.
 XX Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
 KW antinfertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
 KW neuroprotective; antiparkinsonian; protein therapy; EST;
 KW expressed sequence tag; gene; ss.
 OS Homo sapiens.
 XX WO200222660-A2.
 XX 21-MAR-2002.
 XX 10-SEP-2001; 2001WO-US26015.
 XX 11-SEP-2000; 2000US-0659671.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 XX Xue AJ, Yang Y, Wehrman T, Drmanac RT;
 WPI; 2002-292408/33.
 P-PSDB; ABB97309.
 XX An isolated polynucleotide for treating diseases associated with its
 encoded polypeptide such as cancer and multiple sclerosis -
 Claim 1; SEQ ID NO 133; 509pp; English.
 The present invention provides the protein and coding sequences of 444
 novel human proteins. These were isolated from expressed sequences tags
 (ESTs). They can be used to stimulate cell growth, to regulate
 haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 e.g. in burn treatment, to regulate the immune system e.g. to treat
 multiple sclerosis, to regulate activin or inhibin e.g. to treat
 infertility, to regulate haemostasis or thrombolysis e.g. to treat
 stroke and cancer, to screen for drugs, to treat inflammatory conditions
 e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 Parkinson's disease. The present sequence is a coding sequence of the
 invention.

SQ Sequence 514 BP; 138 A; 142 C; 118 G; 116 T; 0 other;

Alignment Scores:
 Pred. No.: 1.62e-17 Length: 514
 Score: 218.50 Matches: 47
 Percent Similarity: 64.89% Conservatives: 14
 Best Local Similarity: 50.00% Mismatches: 32
 Query Match: 42.51% Indels: 1
 DB: 24 Gaps: 1

US-10-045-180A-3 (1-94) x ABNS9722 (1-514)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
 DB 97 ATGAGACCTCGCATCTCTGTCATCTCTGTCGCTGAGCCGAGCCGAGCTGAG 156
 QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAspAsp 39
 DB 157 CCACCTCCAGGCAAGAGCTGATGAGTGTCTGAGCCCGAGAGCATTCGAGGGGACATC 216
 QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
 DB 217 CCAGAAGTGTGTGTTCCCTTGCATGGGACGAAAGCTTGGCTCCAAAGCATCCAGGCTCA 276
 QY 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
 DB 277 AGAATAACATGCGCTGCTGATTCGAGATACAGCGTGCATTCGAGGAGACGTCGCTAT 336
 QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
 DB 337 GGAACCTGCATCTACCGAGGAGACTCTGGGCATCTGCTGC 378

RESULT 11

ABZ11202

ID ABZ11202 standard; cDNA; 652 BP.

AC ABZ11202;

DT 20-JAN-2003 (first entry)

DE Human polynucleotide SEQ ID NO 84.

Human; genome mapping; gene therapy; food supplement; virus; fungus;
 cell-proliferative disorder; neurodegenerative disease; bacterial;
 Parkinson's disease; Alzheimer's disease; autoimmune disease;
 multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 arthritis; cytostatic; immunomodulator; neoplastic; dermatological;
 antiparkinsonian; antidiabetic; immunosuppressive; neuroprotective;
 haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 antiarthritic; gene; ss.

OS Homo sapiens.

PN WO200270539-A2.

PD 12-SEP-2002.

PF 05-MAR-2002; 2002WO-US05095.

PR 05-MAR-2001; 2001US-0799451.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2002-759812/82.

DR P-PSDB; ABP68985.

XX New polynucleotides comprising sequences assembled from expressed
 PT sequence tags (ESTs), useful for treating cell-proliferative,
 PT neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or

PT platelet or coagulation disorders -

PS Claim 1; SEQ ID NO 84; 1012pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising a
 CC nucleotide sequence selected from any of 948 sequences
 CC (ABZ11119-ABZ12066) or their mature protein coding portion, active domain
 CC coding protein or complementary sequences. The polynucleotides are useful
 CC for identifying expressed genes or for physical mapping of human genome.
 CC The encoded polypeptides (ABP68902-ABP69849) are useful as molecular
 CC weight markers, as a food supplement, for generating antibodies, in
 CC medical imaging, screening and diagnostic assays and for treating
 CC cell-proliferative disorders (cancer), neurodegenerative diseases
 CC (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
 CC sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
 CC disorders, platelet or coagulation disorders, wound, burns, incision,
 CC ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
 CC parasitic), arthritis, etc.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 652 BP; 146 A; 185 C; 177 G; 144 T; 0 other;

Alignment Scores:

Pred. No.: 2.21e-17 Length: 652
 Score: 218.50 Matches: 47
 Percent Similarity: 64.89% Conservatives: 14
 Best Local Similarity: 50.00% Mismatches: 32
 Query Match: 42.51% Indels: 1
 DB: 24 Gaps: 1

US-10-045-180A-3 (1-94) x ABZ11202 (1-652)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20

DB 269 ATGAGACCTCGCATCTCTGTCATCTCTGTCGCTGAGCCGAGCCGAGCTGAG 328

QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAspAsp 39

DB 329 CCACCTCCAGGCAAGAGCTGATGAGTGTCTGAGCCCGAGAGCATTCGAGGGACATC 388

QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59

DB 389 CCAGAAGTGTGTGTTCCCTTGCATGGGACGAAAGCTTGGCTCCAAAGCATCCAGGCTCA 448

QY 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79

DB 449 AGGAAAACATGGCTGCTGATTCGAGATACAGCGTGCATTCGAGGAGAGAGCTGCTAT 508

QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93

DB 509 GGAACCTGCATCTACCGAGGAGACTCTGGGCATCTGCTGC 550

RESULT 12

ABZ11183

ID ABZ11183 standard; cDNA; 664 BP.

XX AC ABZ11183;

XX 20-JAN-2003 (first entry)

DE Human polynucleotide SEQ ID NO 65.

Human; genome mapping; gene therapy; food supplement; virus; fungus;
 cell-proliferative disorder; neurodegenerative disease; bacterial;
 Parkinson's disease; Alzheimer's disease; autoimmune disease;
 multiple sclerosis; diabetes; genetic disorder; wound; burn; infection;
 arthritis; cytostatic; immunomodulator; neoplastic; dermatological;
 antiparkinsonian; antidiabetic; immunosuppressive; neuroprotective;
 haemostatic; vulnery; fungicide; antibacterial; virucide; protozoacide;
 antiarthritic; gene; ss.

```

OS Homo sapiens.
XX WO200270539-A2.
XX
XX 12-SEP-2002.
XX
XX 05-MAR-2002; 2002WO-US05095.
XX
XX 05-MAR-2001; 2001US-0799451.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Zhou P, Goodrich RW, Asundi V, Zhang J, Zhao QA, Ren F;
XX Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
XX Wehrman T, Wang J, Wang D, Drmanac RT;
XX
XX WPI; 2002-759812/82.
XX P-PSDB; ABP68966.
XX
XX New polynucleotides comprising sequences assembled from expressed
XX sequence tags (ESTs), useful for treating cell-proliferative,
XX neurodegenerative, autoimmune, genetic, myeloid or lymphoid, or
XX platelet or coagulation disorders -
XX
XX Claim 1; SEQ ID NO 65; 1012pp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising a
XX nucleotide sequence selected from any of 948 sequences
XX (ABZ1119-ABZ12066) or their mature protein coding portion, active domain
XX coding protein or complementary sequences. The polynucleotides are useful
XX for identifying expressed genes or for physical mapping of human genome.
XX The encoded polypeptides (ABP68902-ABP6949) are useful as molecular
XX weight markers, as a food supplement, for generating antibodies, in
XX medical imaging, screening and diagnostic assays and for treating
XX cell-proliferative disorders (cancer), neurodegenerative diseases
XX (Parkinson's or Alzheimer's disease), autoimmune diseases (multiple
XX sclerosis, diabetes, lupus) genetic disorders, myeloid or lymphoid
XX disorders, platelet or coagulation disorders, wound, burns, incision,
XX ulcers, liver or lung fibrosis, infections (bacterial, viral, fungal,
XX parasitic), arthritis, etc.
XX
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 664 BP; 174 A; 168 C; 161 G; 161 T; 0 other;

Alignment Scores:
Pred. No.: 2,266-17 Length: 664
Score: 218.50 Matches: 47
Percent Similarity: 64.89% Conservative: 14
Best Local Similarity: 50.00% Mismatches: 32
Query Match: 42.51% Indels: 1
DB: 24 Gaps: 1

US-10-045-180A-3 (1-94) x ABZ11183 (1-664)

Qy 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
Dy 251 ATGAGGACCTCGCATCTTGTGCTGCCATTCCTCTGTGGCCCTGACGGCCAGGCTGAG 310
Qy 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAsp 39
Dy 311 CCATCCAGGCAAGAGCTGTAGGTTGCTGAGCCGCGGAGCATTTGCGAGCACATC 370
Qy 40 GlnAspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySer 59
Dy 371 CCAGAGAGTGTTGTTCCCTTCATGGGACCAAGACTTGGCTCCAAAGCATCCAGGCTCA 430
Qy 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
Dy 431 AGGAAACACATGGCTGTATTGAGATACACAGGCTGCATTCGAGGAGACGTCGCTAT 490
Qy 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93

```

491 GGACCTGCATCTACCGGGAGAGACTCTGGGCATCTCTCTGC 532

RESULT 13
AAF20911
ID AAF20911 standard; DNA; 464 BP.
XX
AC AAF20911;
XX
DT 14-MAR-2001 (first entry)
XX
XX Human defensin 3 polynucleotide fragment #2478.
XX
XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
XX human; airway disorder; bronchoconstriction; lung inflammation;
XX surfactant depletion; respiratory; bronchodilator; antiinflammatory;
XX immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
XX respiratory obstruction; pulmonary obstruction; impeded respiration;
XX surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
XX respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
XX pulmonary hypertension; emphysema; pulmonary transplantation rejection;
XX chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
XX cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200062736-A2.
XX
XX 26-OCT-2000.
XX
XX 24-MAR-2000; 2000WO-US08020.
XX
XX 06-APR-1999; 99US-0127958.
XX (UYEC-) UNIV EAST CAROLINA.
XX (NYCE/) NYCE J W.
XX
XX Nyce JW;
XX
XX WPI; 2000-679539/66.
XX
XX Low adenosine (A) content antisense oligonucleotides which do not
XX trigger adenosine receptors during metabolism, useful e.g. for treating
XX cancers and respiratory obstructions -
XX
XX Disclosure; Page 218; 1592pp; English.
XX
XX The present invention describes low adenosine (A) content antisense
XX oligonucleotides and compositions (I) comprising them. In the antisense
XX oligonucleotides the A is replaced by a 'Universal' or alternative base.
XX (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
XX immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
XX The antisense oligonucleotides and (I) can be used to down-regulate the
XX expression and or activity of target polypeptides associated with
XX lung/respiratory disorders and malignancies, such as stimulating and
XX activating peptide factors and transmitters, transcription factors,
XX immunoglobulins and antibodies, antibody receptors, cytokines and
XX chemokines, endogenously produced specific and non-specific enzymes,
XX binding proteins, adhesion molecules and their receptors, cytokine and
XX chemokine receptors, adenosine receptors, bradykinin receptors, central
XX nervous system (CNS) and peripheral nervous and non-nervous system
XX receptors, CNS and peripheral nervous and non-nervous system peptide
XX transmitters, defensins, growth factors, vasoactive peptides and
XX receptors, binding proteins and malignancy associated proteins. The
XX antisense oligonucleotides may be used in this way to treat disorders
XX including respiratory obstruction (especially pulmonary obstruction
XX and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
XX and/or surfactant hypoproduction which are associated with a disease or
XX condition selected from pulmonary vasoconstriction, inflammation,
XX allergies, asthma, impeded respiration, respiratory distress syndrome
XX (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
XX hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
XX pulmonary transplantation rejection, pulmonary infections, bronchitis,

CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.
 XX
 SQ Sequence 464 BP; 110 A; 128 C; 119 G; 107 T; 0 other;

Alignment Scores:
 Pred. No.: 2,49e-17 Length: 464
 Score: 216.50 Matches: 47
 Percent Similarity: 64.89% Conservative: 14
 Best Local Similarity: 50.00% Mismatches: 32
 Query Match: 42.12% Indels: 1
 DB: 21 Gaps: 1

US-10-045-180A-3 (1-94) x AAF20911 (1-464)

QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
 Db 90 ATGAGGACCTCGCCATCTTGTGCTGCCATTCCTGTGGCCCTGCAGGCCAGGCTGAG 149

QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAspAsp 39
 Db 150 CCACCTCCAGGCAAGAGCTGATGAGGTGTGTCGACGCCCGAGCAGATTCGACGGGACATC 209

QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
 Db 210 CCAGAAAGTGGTGTTCCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 269

QY 60 ThrLysGlyLeuLeuCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
 Db 270 AGGAAACATGAGCTGCTATTGCAGATACCGAGCGTGCATTCGACGAGGACGCTGCTAT 329

QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
 Db 330 GGAACCTGCATCTACCGAGGAAGACTCTGGGCACTTCTGCTGC 371

RESULT 14
 AAA34789
 ID AAA34789 standard; DNA; 464 BP.
 XX
 AC AAA34789;
 XX
 DT 28-JUL-2000 (first entry)
 XX
 DE Human adenosine receptor related polynucleotide SEQ ID NO:2478.
 XX
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiasthmatic; cytotostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200009525-A2.
 XX
 PD 24-FEB-2000.
 XX
 PF 03-AUG-1999; 99WO-US17712.
 XX
 PR 03-AUG-1998; 98US-0095212.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 XX
 PI Nyce JW;
 XX
 DR WPI; 2000-205971/18.
 XX
 PT New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,

PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers -
 XX
 PS Disclosure; Page 612; 1343pp; English.
 XX
 CC The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
 CC carcinomas, and cancers which may metastasise to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing the
 CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.

SQ Sequence 464 BP; 110 A; 128 C; 119 G; 107 T; 0 other;

Alignment Scores:
 Pred. No.: 2,49e-17 Length: 464
 Score: 216.50 Matches: 47
 Percent Similarity: 64.89% Conservative: 14
 Best Local Similarity: 50.00% Mismatches: 32
 Query Match: 42.12% Indels: 1
 DB: 21 Gaps: 1

US-10-045-180A-3 (1-94) x AAA34789 (1-464)

QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
 Db 90 ATGAGGACCTCGCCATCTTGTGCTGCCATTCCTGTGGCCCTGCAGGCCAGGCTGAG 149

QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAspAsp 39
 Db 150 CCACCTCCAGGCAAGAGCTGATGAGGTGTGTCGACGCCCGAGCAGATTCGACGGGACATC 209

QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
 Db 210 CCAGAAAGTGGTGTTCCTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 269

QY 60 ThrLysGlyLeuLeuCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
 Db 270 AGGAAACATGAGCTGCTATTGCAGATACCGAGCGTGCATTCGACGAGGACGCTGCTAT 329

QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
 Db 330 GGAACCTGCATCTACCGAGGAAGACTCTGGGCACTTCTGCTGC 371

RESULT 15
 AAF21446
 ID AAF21446 standard; DNA; 644 BP.
 XX
 AC AAF21446;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Human defensin 3 polynucleotide fragment #3013.
 XX

Low adenosine antisense oligonucleotide; phosphorothioate; allergy; human; airway disorder; bronchoconstriction; lung inflammation; surfactant depletion; respiratory bronchodilator; antiinflammatory; immunosuppressive; antihistaminic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RUS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

Homo sapiens.

WO200062736-A2.

26-OCT-2000.

24-MAR-2000; 2000WO-US08020.

06-APR-1999; 99US-0127958.

(UYEC-) UNIV EAST CAROLINA.

(NYCE//) NYCE J W.

Nyce JW;

WPI; 2000-679539/66.

Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -

Disclosure; Page 217-218; 1592pp; English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

Sequence 644 BP; 110 A; 182 C; 184 G; 140 T; 28 other;

Alignment Scores:

Pred. No.:	3,838-17	Length:	644
Score:	216.50	Matches:	47
Percent Similarity:	64.89%	Conservative:	14
Best Local Similarity:	50.00%	Mismatches:	32
Query Match:	42.12%	Indels:	1
DB:	21	Gaps:	1

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 15:06:18 ; Search time 325 Seconds
(without alignments)
963.776 Million cell updates/sec

Title: US-10-045-180A-3
Perfect score: 514
Sequence: 1 MRTLTLASFLIALVALQAAE.....GEHLGGTCFILGERYPICCY 94

Scoring table: BLOSUM62

Gapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2211978 seqs, 1666101734 residues

Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10045180/runat_17122003_145026_12635/app_query.fasta_1.860
-DB=Published Applications NA -OFMT=fasta -SUFFIX=rnpb -MINMATCH=0.1
-LOOPLC=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62
-TRANS=human40.cdi -LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10045180 @CEN 1 1 174 @runat_17122003_145026_12635
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORE=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	514	100.0	453	14	US-10-013-770-2	Sequence 2, Appli
2	514	100.0	453	14	US-10-045-180A-2	Sequence 2, Appli
3	393.5	76.6	4415	14	US-10-013-770-1	Sequence 1, Appli
4	393.5	76.6	4415	14	US-10-045-180A-1	Sequence 1, Appli
5	227	44.2	542	14	US-10-045-180A-8	Sequence 473, App
6	227	44.2	860	14	US-10-044-090-473	Sequence 223, App
7	218.5	42.5	555	13	US-10-252-157-223	Sequence 8, Appli
8	216.5	42.1	726	10	US-09-816-828-16	Sequence 16, Appli
9	216.5	42.1	1348	15	US-10-125-237-78	Sequence 78, Appli
10	216.5	42.1	1348	15	US-10-105-891-78	Sequence 11, Appli
11	175.5	34.1	496	13	US-10-141-645-11	Sequence 14, Appli
12	165.5	32.2	500	13	US-10-141-645-14	Sequence 13, Appli
13	165.5	32.2	500	13	US-10-313-994-13	Sequence 1047, Ap
14	160	31.1	412	10	US-09-736-457-1047	Sequence 1047, Ap
15	160	31.1	412	10	US-09-902-941-1047	Sequence 1047, Ap
16	160	31.1	412	10	US-09-849-626-1047	Sequence 1047, Ap
17	160	31.1	412	13	US-10-113-872-1047	Sequence 1047, Ap
18	160	31.1	412	13	US-10-017-754-1047	Sequence 1047, Ap
19	159.5	31.0	353	11	US-09-918-995-18323	Sequence 18323, A
20	157.5	30.6	337	11	US-09-918-995-18138	Sequence 18138, A
21	157.5	30.6	337	11	US-09-918-995-19070	Sequence 19070, A
22	157.5	30.6	495	13	US-10-141-645-16	Sequence 16, Appli
23	157.5	30.6	495	13	US-10-313-994-15	Sequence 15, Appli
24	155	30.2	437	13	US-10-141-645-122	Sequence 18938, A
25	153.5	29.9	372	11	US-09-918-995-18938	Sequence 124, App
26	153	29.8	437	13	US-10-141-645-124	Sequence 120, App
27	152	29.6	437	13	US-10-141-645-120	Sequence 141529, A
28	147	28.6	654	13	US-10-027-632-141529	Sequence 7, Appli
29	147	28.6	654	14	US-10-045-180A-7	Sequence 28, Appli
30	147	28.6	4295	14	US-10-027-632-141529	Sequence 141530, A
31	145.5	28.3	243	13	US-10-313-994-28	Sequence 30054, A
32	145	28.2	741	13	US-10-027-632-141528	Sequence 30054, A
33	145	28.2	741	13	US-10-027-632-141530	Sequence 15342, A
34	145	28.2	741	14	US-10-027-632-141530	Sequence 48280, A
35	145	28.2	741	14	US-10-027-632-141530	Sequence 48281, A
36	140.5	27.3	592	13	US-10-027-632-30053	Sequence 48282, A
37	140.5	27.3	592	13	US-10-027-632-30053	Sequence 48283, A
38	140.5	27.3	592	14	US-10-027-632-30053	Sequence 48284, A
39	140.5	27.3	592	14	US-10-027-632-30054	
40	129.5	25.2	597	9	US-09-864-761-15342	
41	128.5	25.0	598	13	US-10-027-632-48280	
42	128.5	25.0	598	13	US-10-027-632-48281	
43	128.5	25.0	598	13	US-10-027-632-48282	
44	128.5	25.0	598	13	US-10-027-632-48283	
45	128.5	25.0	598	13	US-10-027-632-48284	

ALIGNMENTS

RESULT 1
US-10-013-770-2
; Sequence 2, Application US/10013770
; Publication No. US20020115151A1
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; THERAPEUTIC APPLICATIONS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson & Bear

STREET: 550 West C Street

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/013,770

```

RESULT 2
US-10-045-180A-2
; Sequence 2, Application US/10045180A
; Publication No. US20020182703A1
; GENERAL INFORMATION:
; APPLICANT: Bouguetelret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Composit
; TITLE OF INVENTION: Containing Them and Applications to Diagnosis and to Therapeutic
; FILE REFERENCE: GEN-10001
; CURRENT APPLICATION NUMBER: US/10/045,180A
; CURRENT FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: US 09/486,580
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: PCT/FR98/01864
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: FR 97/10823
; PRIOR FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 453
; TYPE: DNA

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 550 West C Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30 (SP01)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/013,770
FILING DATE: 10-Dec-2001

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/486,580
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.064C1

INFORMATION FOR SEO ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4415 BASE PAIRS
TYPE: NUCLEOTIDE


```
LOCATION: (1150)..(1150)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: CAAT_signal
LOCATION: (1711)..(1714)
OTHER INFORMATION:
FEATURE:
NAME/KEY: TATA_signal
LOCATION: (1758)..(1767)
OTHER INFORMATION:
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1780)..(1780)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1836)..(1874)
OTHER INFORMATION: Exon 1
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1875)..(1880)
OTHER INFORMATION: splice donor site
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1974)..(1974)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2117)..(2117)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2133)..(2133)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2155)..(2155)
OTHER INFORMATION: Alu insertion
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2186)..(2186)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2191)..(2191)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2367)..(2367)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (2710)..(2780)
OTHER INFORMATION: L1 fragment insertion
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3391)..(3393)
OTHER INFORMATION: splice acceptor site
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3394)..(3577)
OTHER INFORMATION: Exon 2
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3406)..(3408)
OTHER INFORMATION: Translation initiation codon (ATG)
FEATURE:
NAME/KEY: misc_feature
LOCATION: (3578)..(3583)
OTHER INFORMATION: splice donor site
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4123)..(4123)

OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4161)..(4163)
OTHER INFORMATION: splice acceptor site
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4164)..(4379)
OTHER INFORMATION: Exon 3
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4274)..(4276)
OTHER INFORMATION: Translation termination codon (TAA)
FEATURE:
NAME/KEY: polyA_signal
LOCATION: (4374)..(4379)
OTHER INFORMATION:
US-10-045-180A-1

Alignment Scores:
Pred. No.: 8,82e-46 Length: 4415
Score: 393.50 Matches: 94
Percent Similarity: 32.41% Conservative: 0
Best Local Similarity: 32.41% Mismatches: 196
Query Match: 76.56% Indels: 196
DB: 14 Gaps: 1

US-10-045-180A-3 (1-94) x US-10-045-180A-1 (1-4415)

QY 1 MetArgThrLeuThrIleuLeuSerAlaPheLeuIleuValAlaLeuGlnAlaTPAlaGln 20
Db 3406 ATGAGGACCCCTCAACCCCTCTGCTGCTTCTCTGCTGAGCCCTTCAAGGCTGGGAGAG 3465
QY 21 ProluGlnAlaArgAlaHisGluMetProAlaGluIleGlnProProAlaAspArgGln 40
Db 3466 CCGCTCCAGGCGAAGAGCTCATGAGATGCCAGCCAGAGAGAGAGAGAGAGATGACAG 3525
QY 41 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGln 55
Db 3526 GATGTGTCTATTACTTTTCAAGAGATGACAGCTGCTCTTCAAGTTCAGGTGAGAGAGA 3585
QY 55 ----- 55
Db 3586 TGCCAGCATCAGAGCTACAGACTACAGACAGAGAGAGAGAGAGCTTGAAATTGGAT 3645
QY 55 ----- 55
Db 3646 CTCAGTGGCAGATGTCATTAGTGGCTATACATCTGTGGTCTGATTTCTCA 3705
QY 55 ----- 55
Db 3706 TATCTAAATGGAATAGAGAACCAAGAAATCTAAGATTTTCTTCCAAAACCTTG 3765
QY 55 ----- 55
Db 3766 ATTCCAAGATATGACTGTGAATTTCACTAGATTTAAGATATAAGAGATGCTACTAGTT 3825
QY 55 ----- 55
Db 3826 CTTTGGAGCCAGACAAACAAGCTTAAGTATATAGAAAAATATTCAACCTGTCTATAT 3885
QY 55 ----- 55
Db 3886 AGGAGTTTAAACCTGAGAGAGAGAGCTAAGATGTGTTCAAGTGTGTGTGATGGCC 3945
QY 55 ----- 55
Db 3946 AGGAATGACAGAAAAGTGAAGCAAGAGAGATGATCTCGAATCCTGTGTGACAGCACTG 4005
QY 55 ----- 55
Db 4006 CTCTGTGATTTATTCTCATTTAGTACTGAGATTGTTGTGTGACCGGCTGAATACAGCCAA 4065
```

QY 55 ----- 55
Db 4066 CATCACTCATTACGCAACATGACTTCTCCAGATTCCCTTACCAACCACTGCTGAC 4125
QY 56 ----- Val-ProglYSerThrlYsgLYleuLYleCY 65
Db 4126 CCCGACTACAGTTTCTGATGCTCTCTCGGCTCCCGAGGCTCAACAAAGGCTTGATCTG 4185
QY 65 SHISCYARValleuTYrCYsIlePheGLYluHISLeuGLYGLYThrCYsPheIlele 85
Db 4186 CCATTGACAGACTACTACTGATTTTGGAGAAACATCTTGTTGGAGGACCTGCTTCATCT 4245
QY 85 uGLYluARgTYrProLYleCYsCYsTYr 94
Db 4246 TGGTGAACGCTACCAATCTGCTGCTAC 4273
RESULT 5
US-10-045-180a-8
; Sequence 8, Application US/10045180A
; Publication No. US20020182703A1
; GENERAL INFORMATION:
; APPLICANT: Bougueteloret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Human Defensein Polypeptide Def-X, Genomic DNA and cDNA, Compositi
; TITLE OF INVENTION: Containing Them and Applications to Diagnosis and to Therapeutic
; FILE REFERENCE: GEN-1001
; CURRENT APPLICATION NUMBER: US/10/045,180A
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: US 09/486,580
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: PCT/FR98/01864
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: FR 97/10823
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 542
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)..(345)
; OTHER INFORMATION: Def-4 (HNP-4) coding sequence
US-10-045-180a-8
Alignment Scores:
Pred. No.: 4,93e-23 Length: 542
Score: 227.00 Matches: 46
Percent Similarity: 64.52% Conservative: 14
Best Local Similarity: 49.46% Mismatches: 33
Query Match: 44.16% Indels: 0
DB: 14 Gaps: 0
US-10-045-180a-3 (1-94) x US-10-045-180a-8 (1-542)
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db 52 ATGAGGATTATCGCCCTCTCGCTGCTATCTTGTGTAGACCTCCAGATCCGGGAGGC 111
QY 21 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLYleGlnProAlaAspArgin 40
Db 112 CCACCTCCAGGCAAGAGGTGATGAGCTCCAGGCAAGAGAGGTGGGCAAGAACAG 171
QY 41 AspValValIleTYrPheSerGLYAspAspSerCYsSerLeuGlnValProGLYSerThr 60
Db 172 GACATATCTATTCTTCTTGATGAGTAAAGCTCTGCTTTCAGGTTTCAGGCTCAACA 231
QY 61 LysGLYleuIleCYsHisCYsArgValLeuTYrCYsIlePheGLYluHISLeuGLYGLY 80
Db 232 AGGGGATGCTGCTGCTTGCATGATTAAGATTCTGCGGGAACAACAATCTGTTGGG 291
QY 81 ThrCYsPheIleLeuGLYluARgTYrProLYleCYsCYs 93

Db 292 AACTGCTCATTTGGTGTGTGAGTTTCACATACCTGCTGC 330
RESULT 6
US-10-044-090-473
; Sequence 473, Application US/10044090
; Publication No. US20020137081A1
; GENERAL INFORMATION:
; APPLICANT: Olga Bandman
; TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
; FILE REFERENCE: PA-0028 US
; CURRENT APPLICATION NUMBER: US/10/044,090
; PRIOR FILING DATE: 2002-01-09
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: PERL Program
; SEQ ID NO 473
; LENGTH: 860
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020137081A1 1670142CB1
US-10-044-090-473
Alignment Scores:
Pred. No.: 9.16e-23 Length: 860
Score: 227.00 Matches: 46
Percent Similarity: 64.52% Conservative: 14
Best Local Similarity: 49.46% Mismatches: 33
Query Match: 44.16% Indels: 0
DB: 14 Gaps: 0
US-10-045-180a-3 (1-94) x US-10-044-090-473 (1-860)
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
Db 76 ATGAGGATTATCGCCCTCTCGCTGCTATCTTGTGTAGACCTCCAGATCCGGGAGGC 135
QY 21 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLYleGlnProAlaAspArgin 40
Db 136 CCACCTCCAGGCAAGAGGTGATGAGCTCCAGGCAAGAGAGGTGGGCAAGAACAG 195
QY 41 AspValValIleTYrPheSerGLYAspAspSerCYsSerLeuGlnValProGLYSerThr 60
Db 196 GACATATCTATTCTTCTTGATGAGTAAAGCTCTGCTTTCAGGTTTCAGGCTCAACA 255
QY 61 LysGLYleuIleCYsHisCYsArgValLeuTYrCYsIlePheGLYluHISLeuGLYGLY 80
Db 256 AGGGGATGCTGCTGCTTGCATGATTAAGATTCTGCGGGAACAACAATCTGTTGGG 315
QY 81 ThrCYsPheIleLeuGLYluARgTYrProLYleCYsCYs 93
Db 316 AACTGCTCATTTGGTGTGTGAGTTTCACATACCTGCTGC 354
RESULT 7
US-10-252-157-223
; Sequence 223, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Fatis, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 223
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens

```
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 903338.12
US-10-252-157-223
```

Alignment Scores:

Pred. No.:	8,51e-22	Length:	555
Score:	218.50	Matches:	47
Percent Similarity:	64.89%	Conservative:	14
Best Local Similarity:	50.00%	Mismatches:	32
Query Match:	42.51%	Indels:	1
DB:	13	Gaps:	1

US-10-045-180a-3 (1-94) x US-10-252-157-223 (1-555)

```
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTTPAlaGlu 20
    |||||
DB 145 ATGAGAGCCCTCGCCATCTTGCTCCATCTCTCGTGGCCCTGACAGCCCGACGCTGAG 204
    |||||
QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAspAsp 39
    |||||
DB 205 CCATCTCCAGGAGAGAGCTGATGAGGTTCCTGACGCCCGGAGCAGATTGCGAGCGGACATC 264
    |||||
QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
    |||||
DB 265 CCAGAGGTGTGTGTTCCCTTGATGGAGGAGAAAGCTTGCTCCAAAGCATTCGAGGCTCA 324
    |||||
QY 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGlnHisLeuGly 79
    |||||
DB 325 AGGAAAAACATGAGCTGCTATGACAAATACCAAGCGTGCATTGACGAGAAAGTCGCTAT 384
    |||||
QY 80 GlyThrCysPheIleLeuGlyGlnArgTyrProIleCysCys 93
    |||||
DB 385 GGAACCTGCATCTACACAGGAAAGACTCTGGGCACTTCCTGCG 426
    |||||
```

RESULT 8

```
US-09-816-828-16
; Sequence 16, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
```

```
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Ping
APPLICANT: Goodrich, Ryle
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aidong J.
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Zhiwei
APPLICANT: Zhao, Qing A.
APPLICANT: Zhang, Jie
APPLICANT: Wang, Jian-Rui
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
FILE REFERENCE: 791CIP2E
CURRENT APPLICATION NUMBER: US/09/816,828
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/770,160
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 09/552,929
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: pt_fl_genes Version 2.0
SEQ ID NO 16
LENGTH: 726
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: CDS
```

```
LOCATION: (151)..(522)
US-09-816-828-16
```

Alignment Scores:

Pred. No.:	2.37e-21	Length:	726
Score:	216.50 <td>Matches:</td> <td>47</td>	Matches:	47
Percent Similarity:	64.89% <td>Conservative:</td> <td>14</td>	Conservative:	14
Best Local Similarity:	50.00% <td>Mismatches:</td> <td>32</td>	Mismatches:	32
Query Match:	42.12% <td>Indels:</td> <td>1</td>	Indels:	1
DB:	10 <td>Gaps:</td> <td>1</td>	Gaps:	1

US-10-045-180a-3 (1-94) x US-09-816-828-16 (1-726)

```
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTTPAlaGlu 20
    |||||
DB 241 ATGAGAGCCCTCGCCATCTTGCTCCATCTCTCGTGGCCCTGACAGCCCGACGCTGAG 300
    |||||
QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProProAlaAspAsp 39
    |||||
DB 301 CCATCTCCAGGAGAGCTGATGAGGTTCCTGACGCCCGGAGCAGATTGCGAGCGGACATC 360
    |||||
QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
    |||||
DB 361 CCAGAGGTGTGTGTTCCCTTGATGGAGGAGAAAGCTTGCTCCAAAGCATTCGAGGCTCA 420
    |||||
QY 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGlnHisLeuGly 79
    |||||
DB 421 AGGAAAAACATGAGCTGCTATGACAAATACCAAGCGTGCATTGACGAGAAAGTCGCTAT 480
    |||||
QY 80 GlyThrCysPheIleLeuGlyGlnArgTyrProIleCysCys 93
    |||||
DB 481 GGAACCTGCATCTACACAGGAAAGACTCTGGGCACTTCCTGCG 522
    |||||
```

RESULT 9

```
US-10-125-237-78
; Sequence 78, Application US/10125237
; Publication No. US20030022329A1
; GENERAL INFORMATION:
```

```
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ping
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Xue, Aidong J.
APPLICANT: Zhang, Jie
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20030022329A1el Nucleic Acids and
FILE REFERENCE: 791CIP2ADIV
CURRENT APPLICATION NUMBER: US/10/125,237
CURRENT FILING DATE: 2002-04-17
PRIOR APPLICATION NUMBER: 09/668,317
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/552,929
PRIOR FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 91
SOFTWARE: pt_fl_genes Version 2.0
SEQ ID NO 78
LENGTH: 1348
TYPE: DNA
ORGANISM: Homo sapiens
```

```
FEATURE:
NAME/KEY: CDS
```

```
LOCATION: (1)..(1143)
US-10-125-237-78
```

Alignment Scores:

Pred. No.:	5.45e-21	Length:	1348
Score:	216.50 <td>Matches:</td> <td>47</td>	Matches:	47
Percent Similarity:	64.89% <td>Conservative:</td> <td>14</td>	Conservative:	14
Best Local Similarity:	50.00% <td>Mismatches:</td> <td>32</td>	Mismatches:	32
Query Match:	42.12% <td>Indels:</td> <td>1</td>	Indels:	1
DB:	15 <td>Gaps:</td> <td>1</td>	Gaps:	1

US-10-045-180A-3 (1-94) x US-10-125-237-78 (1-1348)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
DB 859 ATGAGACCTTCGCGCATCTTCCTGCGTGGCGCTGACGCCGACGCTGAG 918
QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProAlaAsp 39
DB 919 CCATCCGAGGAGAGCTGATGAGTTCCTGACGCGCCGACGATTCGACGCGACATC 978
QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
DB 979 CCAAGAGTGTGTTTCCCTTCGATGAGGACCAAGCTTGCTCCAAAGCATTCGAGGCTCA 1038
QY 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
DB 1039 AGGAAAAACATGACTGCTATTGCGAATACCAAGCTGCATTCGACGAGAAACGTGCTAT 1098
QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
DB 1099 GGAACCTGCATCTTACCAAGGAAAGCTCGGCATTCGCTGC 1140

RESULT 10

US-10-105-891-78

Sequence 78, Application US/10105891
Publication No. US20030073099A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Zhou, Ping
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyan
APPLICANT: Zhao, Qing A.
APPLICANT: Xue, Aidong J.
APPLICANT: Zhang, Jie
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Kui
APPLICANT: Drmanac, Radje T.
TITLE OF INVENTION: No. US20030073099A1el Nucleic Acids and
FILE REFERENCE: 791CIP2A
CURRENT APPLICATION NUMBER: US/10/105,891
PRIORITY FILING DATE: 2002-03-25
PRIORITY APPLICATION NUMBER: 09/668,317
PRIORITY FILING DATE: 2000-09-22
PRIORITY APPLICATION NUMBER: 09/552,929
PRIORITY FILING DATE: 2000-04-18
NUMBER OF SEQ ID NOS: 91
SOFTWARE: pf_fl_genes Version 2.0
SEQ ID NO 78
LENGTH: 1348
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1) .. (1143)
US-10-105-891-78

Alignment Scores:

Pred. No.:	5,45e-21	Length:	1348
Score:	216.50	Matches:	47
Percent Similarity:	64.89%	Conservative:	14
Best Local Similarity:	50.00%	Mismatches:	32
Query Match:	42.12%	Indels:	1
DB:	15	Gaps:	1

US-10-045-180A-3 (1-94) x US-10-105-891-78 (1-1348)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAlaGlu 20
DB 859 ATGAGACCTTCGCGCATCTTCCTGCGTGGCGCTGACGCCGACGCTGAG 918
QY 21 ProLeuGlnAlaArgAlaHisGluMet---ProAlaGlnLysGlnProAlaAsp 39

DB 919 CCATCCGAGGAGAGCTGATGAGTTCGTCGACCCCGGACGATTCGACGCGGACATC 978
QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
DB 979 CCAAGAGTGTGTTTCCCTTCGATGAGGACCAAGCTTGCTCCAAAGCATTCGAGGCTCA 1038
QY 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
DB 1039 AGGAAAAACATGACTGCTATTGCGAATACCAAGCTGCATTCGACGAGAAACGTGCTAT 1098
QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93
DB 1099 GGAACCTGCATCTTACCAAGGAAAGCTCGGCATTCGCTGC 1140

RESULT 11

US-10-141-645-11

Sequence 11, Application US/10141645
Publication No. US20030144184A1
GENERAL INFORMATION:
APPLICANT: Robert Lehner
APPLICANT: Alan Waring
APPLICANT: Alexander Cole
APPLICANT: Teresa Hong
TITLE OF INVENTION: Retrocyclins - Antiviral and
FILE REFERENCE: UCLA-001CIP
CURRENT APPLICATION NUMBER: US/10/141,645
PRIORITY FILING DATE: 2002-05-06
PRIORITY APPLICATION NUMBER: 60/284,855
PRIORITY FILING DATE: 2001-04-18
PRIORITY APPLICATION NUMBER: Unassigned
PRIORITY FILING DATE: 2002-04-18
NUMBER OF SEQ ID NOS: 125
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 496
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (124) ... (304)
OTHER INFORMATION: retrocyclin
US-10-141-645-11

Alignment Scores:

Pred. No.:	1.14e-15	Length:	496
Score:	175.50	Matches:	44
Percent Similarity:	58.51%	Conservative:	11
Best Local Similarity:	46.81%	Mismatches:	38
Query Match:	34.14%	Indels:	1
DB:	13	Gaps:	1

US-10-045-180A-3 (1-94) x US-10-141-645-11 (1-496)

QY 1 MetArgThrLeuThrLeuSerAla---PheLeuValAlaLeuGlnAlaTrpAla 19
DB 76 ATGAGACCTTCGCGCATCTTCCTGCGTGGCGCTGACGCCGACGCTGAGGCTG 135
QY 20 GlnProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProAlaAsp 39
DB 136 GAGCCACTTCAGGCAAGAGCTGATGAGCTGCAGGCCGACGAGAGGCTGAGAGCATGAT 195
QY 40 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59
DB 196 CAGGAATGCTTCATGCTTACATGACATGAAAGTGCCTTCCTCCGCTTCAGACTCA 255
QY 60 ThrLysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGly 79
DB 256 GCGAGAGCTTGAAGTGCATTTGCGGAAAGAGATTTGCCGTTTGAATACGTCGCTTT 315
QY 80 GlyThrCysPheIleLeuGlyGluArgTyrProIleCysCys 93

Db 316 GGGTCTGCGCCTTTGCTGATACCTCACCGGATCTGCTGC 357

RESULT 12

US-10-141-645-14

Sequence 14, Application US/10141645

Publication No. US2003014416A1

GENERAL INFORMATION:

APPLICANT: Robert Lehrer

APPLICANT: Alan Maring

APPLICANT: Alexander Cole

APPLICANT: Teresa Hong

TITLE OF INVENTION: Retrocyclins - Antiviral and

TITLE OF INVENTION: Antimicrobial peptides

FILE REFERENCE: UCLA-001CIP

CURRENT APPLICATION NUMBER: US/10/141,645

PRIOR FILING DATE: 2002-05-06

PRIOR APPLICATION NUMBER: 60/284,855

PRIOR FILING DATE: 2001-04-18

PRIOR APPLICATION NUMBER: Unassigned

NUMBER OF SEQ ID NOS: 125

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 14

LENGTH: 500

TYPE: DNA

ORGANISM: Macaca mulatta

FEATURE:

NAME/KEY: CDS

LOCATION: (95)...(325)

OTHER INFORMATION: theta defensin 1A precursor

FEATURE:

NAME/KEY: sig peptide

LOCATION: (95)...(154)

FEATURE:

NAME/KEY: mat peptide

LOCATION: (287)...(313)

OTHER INFORMATION: ligated to RTD1b in head-to-tail orientation to

OTHER INFORMATION: form the cyclic octadecapeptide RTD1; RTD1 is

OTHER INFORMATION: stabilized by three intramolecular disulfides

US-10-141-645-14

Alignment Scores:

Pred. No.:	3,16e-14	Length:	500
Score:	165.50	Matches:	45
Percent Similarity:	56.57%	Conservative:	11
Best Local Similarity:	45.45%	Mismatches:	32
Query Match:	32.20%	Indels:	11
DB:	13	Gaps:	4

US-10-045-180A-3 (1-94) x US-10-141-645-14 (1-500)

QY 1 MetArgThrLeuThrLeuSerAla---PheLeuLeuValAlaLeuGlnAlaTrpAla 19

Db 95 ATGAGACCTTGCCCTCTCCACCGCATGCTTCTCTGAGCCCTGACCTCAGGCA 154

QY 20 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 39

Db 155 GAGGACGCTCAGGCAAGAGCTGATGAGCTGCGCCACACAGCCTCGAACAAGTAT 214

QY 40 GlnAspValAlaIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59

Db 215 CAGGGATGCGTCATCTCTTACATGCTGAAAGCGCCCTTCCACTTCACAGATCA 274

QY 60 ThrIlyGlyLeuIleCys-----HisCysArgValLeuTyrCysIlePhe 74

Db 275 GCGAAGGCTTGAGGTGATTTGCACACAGAGATTCTCCGTTGTATTAATGT----- 328

QY 75 GlyIuHisLeuGlyGlyThrCysPheIleLeuGlyGlyArgTyrProIleCysCys 93

Db 329 -----CACCTT---GGGTCTGCGCTTTTGTGTTGATCTCACCGGATCTGCTGC 376

RESULT 13

US-10-313-994-13

Sequence 13, Application US/10313994

Publication No. US20030162718A1

GENERAL INFORMATION:

APPLICANT: Selseled, Michael E.

APPLICANT: Tang, Yi-Quan

APPLICANT: Yuan, Jun

APPLICANT: Ouellette, Andre J.

TITLE OF INVENTION: Antimicrobial Thera Defensins and Methods of Using Same

FILE REFERENCE: P-UC 3095

CURRENT APPLICATION NUMBER: US/10/313,994

PRIOR FILING DATE: 2002-12-05

PRIOR APPLICATION NUMBER: US/09/309,487

PRIOR FILING DATE: 1999-05-10

NUMBER OF SEQ ID NOS: 31

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 13

LENGTH: 500

TYPE: DNA

ORGANISM: Macaca mulatta

FEATURE:

NAME/KEY: CDS

LOCATION: (95)...(325)

US-10-313-994-13

Alignment Scores:

Pred. No.:	3,16e-14	Length:	500
Score:	165.50	Matches:	45
Percent Similarity:	56.57%	Conservative:	11
Best Local Similarity:	45.45%	Mismatches:	32
Query Match:	32.20%	Indels:	11
DB:	13	Gaps:	4

US-10-045-180A-3 (1-94) x US-10-313-994-13 (1-500)

QY 1 MetArgThrLeuThrLeuSerAla---PheLeuLeuValAlaLeuGlnAlaTrpAla 19

Db 95 ATGAGACCTTGCCCTCTCCACCGCATGCTTCTCTGAGCCCTGACCTCAGGCA 154

QY 20 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 39

Db 155 GAGGACGCTCAGGCAAGAGCTGATGAGCTGCGCCACACAGCCTCGAACAAGTAT 214

QY 40 GlnAspValAlaIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 59

Db 215 CAGGGATGCGTCATCTCTTACATGAGCTGAAAGCGCCCTTCCACTTCACAGATCA 274

QY 60 ThrIlyGlyLeuIleCys-----HisCysArgValLeuTyrCysIlePhe 74

Db 275 GCGAAGGCTTGAGGTGATTTGCACACAGAGATTCTCCGTTGTATTAATGT----- 328

QY 75 GlyIuHisLeuGlyGlyThrCysPheIleLeuGlyGlyArgTyrProIleCysCys 93

Db 329 -----CACCTT---GGGTCTGCGCTTTTGTGTTGATCTCACCGGATCTGCTGC 376

RESULT 14

US-09-736-457-1047/c

Sequence 1047, Application US/09736457

Parent No. US20020168637A1

GENERAL INFORMATION:

APPLICANT: Wang, Tongtong

APPLICANT: Bangur, Chaitanya S.

APPLICANT: Lodes, Michael A.

APPLICANT: Fanger, Gary

APPLICANT: Vedrick, Tom

APPLICANT: Carter, Darriek

APPLICANT: Retter, Marc

APPLICANT: Mannion, Jane

APPLICANT: Fan, Liqun

APPLICANT: Wang, Aijun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND

FILE REFERENCE: 210121.478C15

CURRENT APPLICATION NUMBER: US/09/736,457

```

; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1047
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1) ... (412)
; OTHER INFORMATION: n = A,T,C or G
US-09-736-457-1047

```

Alignment Scores:

```

Pred. No.: 1.51e-13 Length: 412
Score: 160.00 Matches: 35
Percent Similarity: 53.09% Conservative: 8
Best Local Similarity: 43.21% Mismatches: 38
Query Match: 31.13% Indels: 0
DB: 10 Gaps: 0

```

US-10-045-180A-3 (1-94) x US-09-736-457-1047 (1-412)

```

QY 13 ValAlaLeuGlnAlaTPAlaGluProLeuGlnAlaArgAlaHisGluMetProAlaGln 32
Db 410 GTGGCCCTGAGAGGCCCAAGCTGAGTCACTCCAGAAAGAGCTGATGAGGCTACAAACCCAG 351
QY 33 LysGlnProProAlaAspAspGlnAspValValIleTyrPheSerGlyAspAspSerCys 52
Db 350 AAGCAGTNTGGGGAGAACCAACCAAGACCTTCTCTTCCTTGCGAGGAATGACCTMTNT 291
QY 53 SerLeuGlnValProGlySerThrLysGlyLeuIleCysHisCysArgValLeuTyrCys 72
Db 290 GCTNTTAGAACCTCAGGTTNTCAAGCAAGCAACCTGCTATTGCCAACCAGGCCGCTTGT 231
QY 73 IlePheGlyGluHisLeuGlyGlyThrCysPheIleLeuGlyGluArgTyrProIleCys 92
Db 230 GCTACCCGTAGTCCCTCTCCGGGGTGTGTGAATCAGTGGCCGCTNTACAGACTCTGC 171
QY 93 Cys 93
Db 170 TGT 168

```

RESULT 15

```

US-09-902-941-1047/c
; Sequence 1047, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Warnerakis, Margarita
; APPLICANT: Carter, Patrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedrick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.

```

```

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1047
; LENGTH: 412
; TYPE: DNA

```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 183, 271, 287, 292, 294, 343

```

```

; OTHER INFORMATION: n = A,T,C or G
US-09-902-941-1047

```

Alignment Scores:

```

Pred. No.: 1.51e-13 Length: 412
Score: 160.00 Matches: 35
Percent Similarity: 53.09% Conservative: 8
Best Local Similarity: 43.21% Mismatches: 38
Query Match: 31.13% Indels: 0
DB: 10 Gaps: 0

```

US-10-045-180A-3 (1-94) x US-09-902-941-1047 (1-412)

```

QY 13 ValAlaLeuGlnAlaTPAlaGluProLeuGlnAlaArgAlaHisGluMetProAlaGln 32
Db 410 GTGGCCCTGAGAGGCCCAAGCTGAGTCACTCCAGAAAGAGCTGATGAGGCTACAAACCCAG 351
QY 33 LysGlnProProAlaAspAspGlnAspValValIleTyrPheSerGlyAspAspSerCys 52
Db 350 AAGCAGTNTGGGGAGAACCAACCAAGACCTTGTCTCTTTCGAGGAATGACCTMTNT 291
QY 53 SerLeuGlnValProGlySerThrLysGlyLeuIleCysHisCysArgValLeuTyrCys 72
Db 290 GCTNTTAGAACCTCAGGTTNTCAAGCAAGCAACCTGCTATTGCCAACCAGGCCGCTTGT 231
QY 73 IlePheGlyGluHisLeuGlyGlyThrCysPheIleLeuGlyGluArgTyrProIleCys 92
Db 230 GCTACCCGTAGTCCCTCTCCGGGGTGTGTGAATCAGTGGCCGCTNTACAGACTCTGC 171
QY 93 Cys 93
Db 170 TGT 168

```

Search completed: December 17, 2003, 21:35:57
Job time : 330 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 94 Seconds

(without alignments)
441.383 Million cell updates/sec

Title: US-10-045-180A-3

Perfect score: 514
Sequence: 1 NRTLTLSAFLVALQAWAE.....GHHGTCFTIGERRPICCY 94

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10045180/runat.17122003.145024.12553/app.query.fasta_1.860
-DB=Issued Patents NA -QWMT=fastcap -SUFFIX=rti -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pcio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10045180@cgn2_1.193 @runat.17122003.145024.12553 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6C.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/6D.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	514	100.0	453	4 US-09-486-580A-2
2	393.5	76.6	4415	4 US-09-486-580A-1
3	220	42.8	452	1 US-08-158-189-6
4	220	40.3	452	4 US-09-015-643-1198
5	207	40.3	424	1 US-08-158-189-4
6	173.5	33.8	420	1 US-08-486-013-46
7	173.5	33.8	420	2 US-08-482-279-46
8	173.5	33.8	420	2 US-08-342-268-46
9	173.5	33.8	420	2 US-09-015-968-46
10	173.5	33.8	420	4 US-09-397-386-46
11	171.5	33.4	401	1 US-08-486-013-54
12	171.5	33.4	401	2 US-08-482-279-54

13	171.5	33.4	401	2 US-08-342-268-54	Sequence 54, Appl
14	171.5	33.4	401	3 US-09-015-968-54	Sequence 54, Appl
15	171.5	33.4	401	4 US-09-397-386-54	Sequence 54, Appl
16	168.5	32.8	388	1 US-08-486-013-52	Sequence 52, Appl
17	168.5	32.8	388	2 US-08-482-279-52	Sequence 52, Appl
18	168.5	32.8	388	2 US-08-342-268-52	Sequence 52, Appl
19	168.5	32.8	388	3 US-09-015-968-52	Sequence 52, Appl
20	168.5	32.8	388	4 US-09-397-386-52	Sequence 52, Appl
21	166.5	32.4	377	1 US-08-486-013-48	Sequence 48, Appl
22	166.5	32.4	377	2 US-08-482-279-48	Sequence 48, Appl
23	166.5	32.4	377	2 US-08-342-268-48	Sequence 48, Appl
24	166.5	32.4	377	3 US-09-015-968-48	Sequence 48, Appl
25	166.5	32.4	377	4 US-09-397-386-48	Sequence 48, Appl
26	165.5	32.2	422	1 US-08-486-013-41	Sequence 41, Appl
27	165.5	32.2	422	2 US-08-482-279-41	Sequence 41, Appl
28	165.5	32.2	422	2 US-08-342-268-41	Sequence 41, Appl
29	165.5	32.2	422	3 US-09-015-968-41	Sequence 41, Appl
30	165.5	32.2	422	4 US-09-397-386-41	Sequence 41, Appl
31	165.5	32.2	500	4 US-09-300-487-13	Sequence 13, Appl
32	165.5	32.2	500	4 US-09-967-808-13	Sequence 13, Appl
33	164.5	32.0	422	1 US-08-486-013-42	Sequence 42, Appl
34	164.5	32.0	422	2 US-08-482-279-42	Sequence 42, Appl
35	164.5	32.0	422	2 US-08-342-268-42	Sequence 42, Appl
36	164.5	32.0	422	3 US-09-015-968-42	Sequence 42, Appl
37	164.5	32.0	422	4 US-09-397-386-42	Sequence 42, Appl
38	163.5	31.8	422	1 US-08-486-013-40	Sequence 40, Appl
39	163.5	31.8	422	2 US-08-482-279-40	Sequence 40, Appl
40	163.5	31.8	422	2 US-08-342-268-40	Sequence 40, Appl
41	163.5	31.8	422	3 US-09-015-968-40	Sequence 40, Appl
42	163.5	31.8	422	4 US-09-397-386-40	Sequence 40, Appl
43	160	31.1	365	1 US-08-486-013-43	Sequence 43, Appl
44	160	31.1	365	2 US-08-482-279-43	Sequence 43, Appl
45	160	31.1	365	2 US-08-342-268-43	Sequence 43, Appl

ALIGNMENTS

RESULT 1
US-09-486-580A-2
; Sequence 2, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Knobbe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 BASE PAIRS
; TYPE: NUCLEOTIDE
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA

```

; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-486-580A-2

Alignment Scores:
Pred. No.: 1,06e-56 Length: 453
Score: 514.00 Matches: 94
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-045-180A-3 (1-94) x US-09-486-580A-2 (1-453)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrrPalaglu 20
Db 52 ATGAGAGCCCTCACCCTCCTGCTTCTCTCTGCTGAGCCCTTCAAGCCTGGGAGAG 111
QY 21 ProLeuGlnAlaArgAlaHisgluMetProAlaGlnLysGlnProProAlaAspAspGln 40
Db 112 CCGCTCCAGGCAAGAGCTCATGATGACAGCCCAAGACAGCCTCCAGACATGACAG 171
QY 41 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 60
Db 172 GATGGGCTATTACTTTTCAGAGATGAAGAGCTGCTCTCTTCAAGCTTCAAGGCTCAACA 231
QY 61 LysGlyLeuIleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGly 80
Db 232 AAGGGCTTATCTGCACATTCAGAGATGACTACTACTGATTTTGGAGAACATCTGGTGG 291
QY 81 ThrCysPheIleLeuGlyGluArgTyrProIleCysCysTyr 94
Db 292 ACCGTCTCATCTTGTGTGAAGCTACCAATCTGTGCTAC 333

RESULT 2
US-09-486-580A-1
; Sequence 1, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Knobe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4415 BASE PAIRS
; TYPE: NUCLEOTIDE
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: DNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
```

```

; NAME/KEY: Exon 1
; LOCATION: 1836..1874
; FEATURE:
; NAME/KEY: Exon 2
; LOCATION: 3394..3577
; FEATURE:
; NAME/KEY: Exon 3
; LOCATION: 4161..4380
; FEATURE:
; NAME/KEY: start CDS
; LOCATION: 3406..3408
; FEATURE:
; NAME/KEY: stop CDS
; LOCATION: 4276..4278
; FEATURE:
; NAME/KEY: polyadenylation site
; LOCATION: 4374..4379
US-09-486-580A-1

Alignment Scores:
Pred. No.: 4.72e-40 Length: 4415
Score: 393.50 Matches: 94
Percent Similarity: 32.41% Conservative: 0
Best Local Similarity: 32.41% Mismatches: 0
Query Match: 76.56% Indels: 196
DB: Gaps: 1

US-10-045-180A-3 (1-94) x US-09-486-580A-1 (1-4415)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrrPalaglu 20
Db 3406 ATGAGAGCCCTCACCCTCCTGCTTCTCTCTGCTGAGCCCTTCAAGCCTGGGAGAG 3465
QY 21 ProLeuGlnAlaArgAlaHisgluMetProAlaGlnLysGlnProProAlaAspAspGln 40
Db 3466 CCGCTCCAGGCAAGAGCTCATGATGACAGCCCAAGACAGCCTCCAGACATGACAG 3525
QY 41 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGln----- 55
Db 3526 GATGGGCTATTACTTTTCAGAGATGACAGCTGCTCTCTTCAAGTTCCAGTGAAGAGA 3585
QY 55 ----- 55
Db 3586 TGCCAGATGAGAGCTACAGCTAGACAGAGAGACAGAGAGAGCTTGGATTGGAT 3645
QY 55 ----- 55
Db 3646 CTCAGTGCAGATGTCACTAGTGGCTATTACTTAACATCTCTGCTGATTTTCTCA 3705
QY 55 ----- 55
Db 3706 TATCTAATGGAATGAGAACCAAGAAATCTAAGAGATTTTCTTCCAAAAACTTG 3765
QY 55 ----- 55
Db 3766 ATTCAGATATGACTGTGAATTCATAGATTAAAGATATAGAGATGCTACTAGTT 3825
QY 55 ----- 55
Db 3826 CCTTGTGAGCCAGCAAAAGAGCTTAAGTATATAGAAAATATTTCACCTGTCTATAT 3885
QY 55 ----- 55
Db 3886 AGAGGTTTGAACCTGGAGAGAGAGCTTAAGAAATGTGTTCAGGTGTGTATGGGC 3945
QY 55 ----- 55
Db 3946 AGAATGAGAAAAGTGAAGCAAGAGAGATGAGTCCGAATCTGTGTGACACAGACTG 4005
QY 55 ----- 55
Db 4006 CTCTGTATTATTATCTTATTGACTGAGATTGTTGTCTACCGGTGTAAATACAGCAA 4065
```


QY 55 ----- 55
DB 4066 CATCACTATAGCCCAATGACTTCTCCAAAGTTCCTTACCACTGCTGAC 4135
QY 56 -----Val-ProGlySerThrLysGlyLeuIleCY 65
DB 4126 CCCGACTCACTTCTGANTGCTCTCTGCGTCCCAAGCTCAAAAGGCTGATCTG 4185
QY 65 hHicYalrValLeuIleuYrCySilePheGlyGlyIleuGlyGlyThrCySphLeile 85
DB 4186 CCATTGACAGACTACTACTGATTTTGAGAAACATCTGCTGGAACCTGCTCATCT 4245
QY 85 uGlyGlyArgTrpProIleCySyr 94
DB 4246 TGGTAAAGCTAACCAATCTGCTGAC 4273

RESULT 3

US-08-158-189-6
Sequence 6, Application US/08158189
Patent No. 5641497

GENERAL INFORMATION:
APPLICANT: Bevins, Charles L.
APPLICANT: Jones, Douglas E.
TITLE OF INVENTION: Gastrointestinal Defense Peptides
TITLE OF INVENTION: cDNA Sequences, Methods for Production and Use Thereof
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5641497r1s
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,189
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/888,232
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Johnson, Philip S.
REGISTRATION NUMBER: 27,200
REFERENCE/DOCKET NUMBER: CH-0219
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 19..321
US-08-158-189-6

Alignment Scores:

Pred. No.: 2,71e-19 Length: 452
Score: 220.00 Matches: 49
Percent Similarity: 63.00% Conservative: 14
Best Local Similarity: 49.00% Mismatches: 29
Query Match: 42.80% Indels: 8
DB: 1 Gaps: 2

US-10-045-180a-3 (1-94) x US-08-158-189-6 (1-452)

QY 1 MetArgThrLeuThrIleuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
DB 19 ATGGAACCCCTCACCATCTTCACCTGCTCTCTCGGGCCCTCCAGGCCAAGCTGAG 78
QY 21 -----ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLys 33
DB 79 CCACTCCAGCTGAGATGATTCACCTGAGGCAAAAGCTTATGAGGCTGATGCCAGAG 138
QY 34 GlnProAlaAspGlnAspValIleTyrPheSerGlyAspAspSerCysSer 53
DB 139 CAGGTGGGGCAATGACAGAGACTTGGCCCTCTTGGAGAGATGACAGCTCAAGT 198
QY 54 LeuGlnValProGlySerThrLysGlyLeuIleCysHisCySyrValLeuTyrCysIle 73
DB 199 CTTAGAGCTTGGGCTCAACAGAGGCTTTCACCTTGCATTCAGACA---AGGTCTGTAT 255
QY 74 PheGlyGlyHisLeuGlyGlyThrCySphLeileuGlyGlyAlaGlyTrpProIleCySyr 93
DB 256 TCAACAGAAATTCCTATGGGACCTGACCTGCTCATGGGTATTAAACAGATTCGTGTC 315

RESULT 4

US-09-016-434-1198
Sequence 1198, Application US/09016434
Patent No. 6500938

GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERBWITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1198:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: g181546
US-09-016-434-1198

Alignment Scores:

Pred. No.: 2,71e-19 Length: 452
Score: 220.00 Matches: 49
Percent Similarity: 63.00% Conservative: 14
Best Local Similarity: 49.00% Mismatches: 29

```

Query Match: 42.80% Indels: 8
DB: 4 Gaps: 2
US-10-045-180A-3 (1-94) x US-09-016-434-1198 (1-452)
QY 1 MelarqThrluThrluThrluSerAlaPhleluenuValAlenuGlnAlaTrpAlaGlu 20
DB 19 ATGAACCCCTCACCATCTACCTACGCTGTTCTCTCGTGGCCCTCCAGAGCAAGGCTGAG 78
QY 21 -----ProleuGlnAlaArgAlaHisGluMetProAlaGluys 33
DB 79 CCACCTCAAGCTGAGATGATTCACCTCAGGCAAAAGCTTATAGAGCTATGCCAGAG 138
QY 34 GlnProProAlaAspAspGlnAspValValIleTyrPheSerGlyAspAspSerCysSer 53
DB 139 CAGCGTGGGGCAATATGACCAAGACTTTGCCGCTCTCTTGCAGAGAGATCAAGCTCAAGT 198
QY 54 leuGlnValProGlySerThrIlySerGlyLeuIleCysHisCysArgValLeuTyrCysIle 73
DB 199 CTTAGAGCTTTGGCTCTCAACAGAGGCTTTGACCTTGCCATTGGACGA--AGGTCTGTTTAT 255
QY 74 PheGlyGlnHisLeuGlyGlyTyrThyCysPheIleleuGlyGlnArgTyrProIleCysCys 93
DB 256 TCAACAGAAATTTCTCTATGGAGCTGCAGCTGTATAGGATTTAACACAGATTTCTGCTGC 315
RESULT 5
US-08-158-189-4
; Sequence 4, Application US/08158189
; Patent No. 5641497
; GENERAL INFORMATION:
; APPLICANT: Bevin, Charles L.
; APPLICANT: Jones, Douglas E.
; TITLE OF INVENTION: Gastrointestinal Defensin Peptides,
; TITLE OF INVENTION: cDNA Sequences, Methods for Production and Use Thereof
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5641497r18
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,189
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,232
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnson, Philip S.
; REGISTRATION NUMBER: 27,200
; REFERENCE/DOCKET NUMBER: CH-0219
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 424 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 10..294
; US-08-158-189-4
Alignment Scores:

```

```

Pred. No.: 1,12e-17 Length: 424
Score: 207.00 Matches: 44
Percent Similarity: 59.14% Conservative: 11
Best Local Similarity: 47.31% Mismatches: 38
Query Match: 40.27% Idels: 0
DB: 1 Gaps: 0

US-10-045-180A-3 (1-94) x US-08-158-189-4 (1-424)

QY 1 MetArgThrIreuthrIreuthrIreuthrSerAlaPheIleuIleValAlaIleuGlnAlaIleTPRAlaGlu 20
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 10 ATAGAGACATGCCCATCTCTTGTCTGCATCTTCTCTGAGGCTTCTGAGGCTTCTGAGGCTTCTGAG 69
QY 21 ProIleuGlnAlaIleValAlaIleuIleTPRAlaGlnIleValProIleAlaAspAspGln 40
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 70 TCACTCCAGGAAGAGACTGATGAGGCTTCAACCCAGAGAGAGCTGCGGGAGAGAACAGAG 129
QY 41 AspValValIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 60
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 130 GACCTGTCATCTCTCTTCTGAGGAATGACTCTCTGCTTTCAGACCTCAGGTTCTCAG 183
QY 61 IysGlyIleuIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 80
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 190 GCAGAGACCACTGCTATTGAGGCTTCAACCCAGGCGGCTTGTCTACCCGTGATCCCTCTCCAGG 249
QY 81 ThrCysPheIleIleuGlyIleuIleIleIleIleIleIleIleIleIleIleIleIleIleIle 93
   |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 250 GTGTGTGAATCAGTGGCGCGCTTCAAGACCTCTCTCT 288

RESULT 6
US-08-486-013-46
; Sequence 46, Application US/08486013
; Patent No. 5731149
; GENERAL INFORMATION:
; APPLICANT: Seistad, Michael E.
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antibiotic Cryptidin Peptides and Methods
; TITLE OF INVENTION: of Their Use
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,013
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/342,268
; FILING DATE: 18-NOV-1994
; APPLICATION NUMBER: US 07/930,649
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,020
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 1206
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 420 base pairs

```

```

; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-486-013-46

```

```

Alignment Scores:
Pred. No.: 2,03e-13 Length: 420
Score: 173.50 Matches: 38
Percent Similarity: 55.91% Conservative: 14
Best Local Similarity: 40.86% Mismatches: 40
Query Match: 33.75% Indels: 1
DB: 1 Gaps: 1

```

```

US-10-045-180A-3 (1-94) x US-08-486-013-46 (1-420)
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 44 ATGAAGACACTAAATCCCTCTCTCTGCCCCCTGCTCTGCGCTTCCAGGTCACAGCTGAT 103
QY 21 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlysgInProProAlaAspAspGln 40
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 104 CCTATCCAAAACACAGATGAAGACCTAAACTGAGAGACGACGAGGGAAGACGACGAG 163
QY 41 AspValValIleTyrrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 164 GCCGTATCTGCTCTCTTGGAGACCCAGAAAGCTCTTCTTCAA--GAGGAATCGTTG 220
QY 61 LysGlyLeuIleCysHisCysArgValLeuTyrrCysIlePheGlyGlnHisLeuGlyGly 80
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 221 AGAATCTGTATCTCTATTTGTGAACAAGAGCTGCAAAAGAAAGAACATGAATGGG 280
QY 81 ThrCysPheIleLeuGlyGluArgTyrrProIleCysCys 93
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 281 ACCTGCAGAAAGGCTCATTTAATGTACACGCTGCTGCT 319

```

```

RESULT 7
US-08-482-279-46
; Sequence 46, Application US/08482279
; Patent No. 5840498
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,279
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/342,268
; FILING DATE: 18-NOV-1994
; APPLICATION NUMBER: US 07/930,649
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,020
; FILING DATE: 26-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UC 1206

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 420 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-482-279-46

```

```

Alignment Scores:
Pred. No.: 2,03e-13 Length: 420
Score: 173.50 Matches: 38
Percent Similarity: 55.91% Conservative: 14
Best Local Similarity: 40.86% Mismatches: 40
Query Match: 33.75% Indels: 1
DB: 2 Gaps: 1

```

```

US-10-045-180A-3 (1-94) x US-08-482-279-46 (1-420)
QY 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAlaGlu 20
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 44 ATGAAGACACTAAATCCCTCTCTCTGCCCCCTGCTCTGCGCTTCCAGGTCACAGCTGAT 103
QY 21 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlysgInProProAlaAspAspGln 40
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 104 CCTATCCAAAACACAGATGAAGACCTAAACTGAGAGACGACGAGGGAAGACGACGAG 163
QY 41 AspValValIleTyrrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 60
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 164 GCCGTATCTGCTCTTGGAGACCCAGAAAGCTCTTCTTCAA--GAGGAATCGTTG 220
QY 61 LysGlyLeuIleCysHisCysArgValLeuTyrrCysIlePheGlyGlnHisLeuGlyGly 80
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 221 AGAATCTGTATCTATTTGTGAACAAGAGCTGCAAAAGAAAGAACATGAATGGG 280
QY 81 ThrCysPheIleLeuGlyGluArgTyrrProIleCysCys 93
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 281 ACCTGCAGAAAGGCTCATTTAATGTACACGCTGCTGCT 319

```

```

RESULT 8
US-08-342-268-46
; Sequence 46, Application US/08342268
; Patent No. 5844072
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antibiotic Cryptdin Peptides and Methods
; NUMBER OF SEQUENCES: 70
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/342,268
; FILING DATE: 18-NOV-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/930,649
; FILING DATE: 14-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/889,020
; FILING DATE: 26-MAY-1992

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 543.622 Seconds

(Without alignments)
1429.821 Million cell updates/sec

Title: US-10-045-180a-4

Sequence: 1 MRTLTLISAFILVALQAWA 19

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODL=frame+g2n.model -DEV=xlp
-O=/cgn2.1/USPRO.spool_p/US10045180/runat.17122003.145023.12528/app.query.fasta_1.860
-DB=genembl -QWNT=fastcap -SUPFIX=rge -MINMATCH=0 -LLOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCALLIGN=200 -THR.SCOR=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10045180 @CGN 1 1 5283 @runat.17122003.145023.12528 -NCPU=6 -ICPU=3
-NO MAP -LARGOQUERY -NEG.SCOR=0 -WAIT -DSPBLOCK=100 -LONCAL
-DEV.TIMEOUT=120 -MAV.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database:

Genembl:
1: gb_ba:
2: gb_hcg:
3: gb_in:
4: gb_ov:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_scs:
12: gb_sy:
13: gb_un:
14: gb_vl:
15: em_ba:
16: em_fun:
17: em_hum:
18: em_in:
19: em_mn:
20: em_om:
21: em_or:
22: em_ov:
23: em_pat:
24: em_ph:
25: em_pl:
26: em_ro:
27: em_scs:
28: em_un:

29: em_vi:
30: em_hcg_hum:
31: em_hcg_inv:
32: em_hcg_other:
33: em_hcg_mus:
34: em_hcg_pln:
35: em_hcg_rtd:
36: em_hcg_mam:
37: em_hcg_vtc:
38: em_hcg:
39: em_hcg_hum:
40: em_hcg_mus:
41: em_hcg_other:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	90	100.0	453	6 A98571	A98571 Sequence 2
2	90	100.0	453	6 BD074746	BD074746 Human def
3	90	100.0	4415	6 A98570	A98570 Sequence 1
4	90	100.0	4415	6 BD074745	BD074745 Human def
5	85	94.4	128544	9 AF238378	AF238378 Homo sapi
6	85	72.2	232180	2 AC021883	AC021883 Homo sapi
7	64	71.1	398	6 AX405993	AX405993 Sequence
8	64	71.1	452	6 AR270635	AR270635 Sequence
9	64	71.1	452	6 I49588	I49588 Sequence 6
10	64	71.1	452	6 HUMDEFNSIX	HUMDEFNSIX
11	64	71.1	3060	9 HSU33317	HSU33317 Human defen
12	64	71.1	79730	9 AF314060	AF314060 Homo sapi
13	64	71.1	150940	9 AF233439	AF233439 Homo sapi
14	64	71.1	159653	9 AF200455	AF200455 Homo sapi
15	64	71.1	168702	9 AC079018	AC079018 Homo sapi
16	64	71.1	169044	9 AF205406	AF205406 Homo sapi
17	63	70.0	195	6 I49590	I49590 Sequence 9
18	63	70.0	202	6 I49589	I49589 Sequence 8
19	63	70.0	479	4 RABNPGS4A	RABNPGS4A
20	63	70.0	495	4 RABMCP1A	RABMCP1A
21	63	70.0	500	4 S55582	S55582 Oryctolagus
22	63	70.0	532	4 RABMCP2A	RABMCP2A
23	63	70.0	2097	4 RABMCP1AB	RABMCP1AB
24	63	70.0	2133	4 RABMCP1AB	RABMCP1AB
25	61	67.8	110	6 AR021138	AR021138 Sequence
26	61	67.8	110	6 I33694	I33694 Sequence 9
27	61	67.8	110	6 I74645	I74645 Sequence 9
28	61	67.8	110	6 I74691	I74691 Sequence 9
29	61	67.8	203	6 I49591	I49591 Sequence 10
30	61	67.8	203	6 I49592	I49592 Sequence 11
31	61	67.8	218	6 I49596	I49596 Sequence 38
32	61	67.8	347	9 AF188268	AF188268 Macaca mu
33	61	67.8	347	9 AF188269	AF188269 Macaca mu
34	61	67.8	347	9 AF188270	AF188270 Macaca mu
35	61	67.8	433	9 AF184159	AF184159 Macaca mu
36	61	67.8	433	9 AF184160	AF184160 Macaca mu
37	61	67.8	448	9 HUMDEF1A	HUMDEF1A
38	61	67.8	451	9 HUMDEF2A	HUMDEF2A
39	61	67.8	464	9 HSDER3	HSDER3
40	61	67.8	478	6 BD028213	BD028213 Sequence
41	61	67.8	482	9 HSHR1	HSHR1
42	61	67.8	498	6 BD174049	BD174049 Method of
43	61	67.8	498	6 HUMDEF1AA	HUMDEF1AA
44	61	67.8	514	6 AX405718	AX405718 Sequence
45	61	67.8	556	9 BC027917	BC027917 Homo sapi

RESULT 1

ALIGNMENTS

A98571
LOCUS A98571 453 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 2 from Patent WO9911663.
ACCESSION A98571
VERSION A98571.1 GI:6781627
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 453)
AUTHORS Bougueleret, L. and Chumakov, I.
TITLE HUMAN DEFENSIN DEF-X, GENE AND DNAC, COMPOSITION CONTAINING SAME
AND DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
JOURNAL Patent: WO 9911663-A 2 11-MAR-1999;
BOUGUELERET LYDIE (FR); CHUMAKOV ILVA (FR)
FEATURES
source
1..453
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 107 a 121 c 97 g 128 t
ORIGIN

Alignment Scores:
Pred. No.: 3.61e-07 Length: 453
Score: 90.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-045-180A-4 (1-19) x A98571 (1-453)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrrpAla 19
Db 52 ATGAGAGACCCCTCACCCCTCTCTGCTTCTCTGCTGAGGCTTGAGGCTTGAGGCA 108

RESULT 2
LOCUS BD074746 453 bp DNA linear PAT 27-AUG-2002
DEFINITION Human defensin polypeptide Def-X, genome DNA and cDNA, composition
containing the same, and application to diagnosis and remedy
treatment.
ACCESSION BD074746
VERSION BD074746.1 GI:22620349
KEYWORDS JP 2001514264-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Primates; Carnivora; Homiidae; Homo.
1 (bases 1 to 453)
Bougueleret, L. and Shumakov, I.
Human defensin polypeptide Def-X, genome DNA and cDNA, composition
containing the same, and application to diagnosis and remedy
Patent: JP 2001514264-A 2 11-SEP-2001;
JOURNAL
GENE
OS Homo sapiens (human)
PN JP 2001514264-A/2
PD 11-SEP-2001
PF 28-AUG-1998 JP 2000508701
PR 29-AUG-1997 FR 97/10823
PI LYDIE BOUGUELERET, ILVA SHUMAKOV
PC C07K14/435, A01N43/50, A01N63/00, A61K7/00, A61K38/00, A61P23/00,
PC A61P35/00,
PC A61P37/02, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N15/09, PC
C12P21/02,
PC C12P21/08, C12Q1/68, G01N33/53, A61K37/02, C12N15/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Human defensin polypeptide Def-X, genome DNA and cDNA, CC
composition
CC containing the same, and application to diagnosis and remedy
treatment

FEATURES
source
1..453
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 107 a 121 c 97 g 128 t
ORIGIN

Alignment Scores:
Pred. No.: 3.93e-06 Length: 4415
Score: 90.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-045-180A-4 (1-19) x A98570 (1-4415)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrrpAla 19
Db 3406 ATGAGAGACCCCTCACCCCTCTCTGCTTCTCTGCTGAGGCTTGAGGCTTGAGGCA 3462

RESULT 4
LOCUS BD074745 4415 bp DNA linear PAT 27-AUG-2002
DEFINITION Human defensin polypeptide Def-X, genome DNA and cDNA, composition
containing the same, and application to diagnosis and remedy
treatment.
ACCESSION BD074745
VERSION BD074745.1 GI:22620348
KEYWORDS JP 2001514264-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Primates; Carnivora; Homiidae; Homo.
1 (bases 1 to 4415)
Bougueleret, L. and Chumakov, I.
HUMAN DEFENSIN DEF-X, GENE AND DNAC, COMPOSITION CONTAINING SAME
AND DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
Patent: WO 9911663-A 1 11-MAR-1999;
BOUGUELERET LYDIE (FR); CHUMAKOV ILVA (FR)
FEATURES
source
1..4415
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
exon 1836..1874
exon 3394..3577
exon 4161..4380
polyA site 4374..4379
BASE COUNT 1128 a 1109 c 912 g 1252 t 14 others
ORIGIN

Alignment Scores:
Pred. No.: 3.93e-06 Length: 4415
Score: 90.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0

US-10-045-180A-4 (1-19) x A98570 (1-4415)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrrpAla 19
Db 3406 ATGAGAGACCCCTCACCCCTCTCTGCTTCTCTGCTGAGGCTTGAGGCTTGAGGCA 3462

RESULT 4
LOCUS BD074745 4415 bp DNA linear PAT 27-AUG-2002
DEFINITION Human defensin polypeptide Def-X, genome DNA and cDNA, composition
containing the same, and application to diagnosis and remedy
treatment.
ACCESSION BD074745
VERSION BD074745.1 GI:22620348
KEYWORDS JP 2001514264-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Butheria; Primates; Carnivora; Homiidae; Homo.
1 (bases 1 to 4415)
Bougueleret, L. and Chumakov, I.
HUMAN DEFENSIN DEF-X, GENE AND DNAC, COMPOSITION CONTAINING SAME
AND DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
Patent: WO 9911663-A 1 11-MAR-1999;
BOUGUELERET LYDIE (FR); CHUMAKOV ILVA (FR)
FEATURES
source
1..4415
Location/Qualifiers
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
exon 1836..1874
exon 3394..3577
exon 4161..4380
polyA site 4374..4379
BASE COUNT 1128 a 1109 c 912 g 1252 t 14 others
ORIGIN

ACCESSION	BD074745
VERSION	BD074745.1
KEYWORDS	GI:22620348
SOURCE	JP 2001514264-A/1.
ORGANISM	Homo sapiens (human)
	taxid:9606

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo

```

/misc_feature /map="p22-p21"
/misc_feature /clone="SCB-561b17"
/misc_feature 44529..44595
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 45703
/misc_feature /note="low quality region"
/misc_feature 45815..45872
/misc_feature /note="low quality region"
/misc_feature 45824..45872
/misc_feature /note="single stranded/single chemistry region"
unsure 45843
unsure 45847
unsure 45849
unsure 45858
unsure 45868
/misc_feature 46235
/misc_feature /note="low quality region"
/misc_feature 47899..47900
/misc_feature /note="low quality region"
/misc_feature 47899..47900
/misc_feature 49172
/misc_feature /note="low quality region"
unsure 49172
/misc_feature 49220
/misc_feature /note="low quality region"
/misc_feature 49339
/misc_feature /note="low quality region"
/misc_feature 49756..49807
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 49961
/misc_feature /note="low quality region"
/misc_feature 52933..52948
/misc_feature /note="low quality region"
unsure 52943
/misc_feature 53544
/misc_feature /note="low quality region"
/misc_feature 53612..53926
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 53674..53677
/misc_feature /note="low quality region"
/misc_feature 53779..53799
/misc_feature /note="low quality region"
unsure 53787
/misc_feature 54734..55129
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 56028..56165
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 57952..58060
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 57968..57974
/misc_feature /note="low quality region"
unsure 57968
unsure 57970
unsure 57974
/misc_feature 58652
/misc_feature /note="low quality region"
/misc_feature 59561..59638
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 59925
/misc_feature /note="low quality region"
unsure 59925
/misc_feature /note="low quality region"
/misc_feature 61695..62103
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 61858..61859
/misc_feature /note="low quality region"
/misc_feature 62263..62485
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 62655..63118
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 62750..62785
/misc_feature /note="low quality region"

```

```

unsure 62774..62777
unsure 62783
/misc_feature 62801..62825
/misc_feature /note="low quality region"
unsure 62819
unsure 62822
/misc_feature 62872..62874
/misc_feature /note="low quality region"
/misc_feature 62908
/misc_feature /note="low quality region"
/misc_feature 62917..62987
/misc_feature /note="low quality region"
unsure 62933..62935
unsure 62955
unsure 62957
/misc_feature 63082
/misc_feature /note="low quality region"
/misc_feature 63110
/misc_feature /note="low quality region"
/misc_feature 63119..63281
/misc_feature /note="low quality region"
/misc_feature 63119..63141
/misc_feature /note="single clone coverage"
unsure 63129..63131
/misc_feature 63142..63258
/misc_feature /note="single stranded/single chemistry region"
unsure 63159
unsure 63197..63258
unsure 63281
/misc_feature 63876..63878
/misc_feature /note="low quality region"
/misc_feature 63933..64021
/misc_feature /note="single stranded/single chemistry region"
/misc_feature 63943
/misc_feature /note="low quality region"
/misc_feature 63984..63985
/misc_feature /note="low quality region"
/misc_feature 64004..64007
/misc_feature /note="low quality region"
/misc_feature 70360
/misc_feature /note="low quality region"

```

Alignment Scores:

```

Pred. No.: 0.00112 Length: 128544
Score: 85.00 Matches: 18
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.44% Indels: 0
DB: Gaps: 0

```

US-10-045-180A-4 (1-19) x AF238378 (1-128544)

```

QY 2 ArgThrIleuThrlleuSerAla pheIleuValAlaIleuGlnAlaTrpAla 19
Db 63320 AGGACCCCTACCCCTCTCTGCGCTTCTCTCGTGCGCCCTTCAGGCGGCA 63267

```

RESULT 6

AC021883/c 232180 bp DNA linear HTG 25-ANG-2000

LOCUS Homo sapiens chromosome 1 clone RP11-586H15, *** SEQUENCING IN PROGRESS *** 40 unordered pieces.

ACCESSION

AC021883 AC021883.4 GI:9910096

VERSION

HTG: HTGS PHASE1.

KEYWORDS

Homo sapiens (human)

SOURCE

ORGANISM

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 232180)

AUTHORS Waterston, R.H.

TITLE The sequence of Homo sapiens clone

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 232180)

AUTHORS
TITLE
JOURNAL

COMMENT

Waterston, R.H.
Direct Submission
Submitted (21-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Aug 25, 2000 this sequence version replaced gi:9838290.

----- Genome Center -----

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>

----- Project Information -----

Center project name: H.NH0588H15

----- Summary Statistics -----

Sequencing vector: M13; 86%
Sequencing vector: plasmid; 14%
Chemistry: Dye-terminator Big Dye; 14% of reads
Chemistry: Dye-terminator Big Dye; 14% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 211766 bases at least Q40
Consensus quality: 217990 bases at least Q30
Consensus quality: 221643 bases at least Q20
Insert size: 181000; agarose-fp
Insert size: 228280; sum-of-contigs
Quality coverage: 6.35 in Q20 bases; agarose-fp
Quality coverage: 4.69 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1      1312: contig of 1312 bp in length
*      1313      1412: gap of unknown length
*      1413      2446: contig of 1034 bp in length
*      2447      2546: gap of unknown length
*      2547      3878: contig of 1332 bp in length
*      3879      3978: gap of unknown length
*      3979      5215: contig of 1237 bp in length
*      5216      5315: gap of unknown length
*      5316      6544: contig of 1229 bp in length
*      6545      6644: gap of unknown length
*      6644      7940: contig of 1296 bp in length
*      7941      8040: gap of unknown length
*      8041      9190: contig of 1150 bp in length
*      9191      9290: gap of unknown length
*      9291      10909: contig of 1619 bp in length
*      10910      11009: gap of unknown length
*      11010      12203: contig of 1194 bp in length
*      12204      12303: gap of unknown length
*      12304      13542: contig of 1239 bp in length
*      13543      13642: gap of unknown length
*      13643      14982: contig of 1340 bp in length
*      14983      15082: gap of unknown length
*      15083      16905: contig of 1823 bp in length
*      16906      17005: gap of unknown length
*      17006      18051: contig of 1046 bp in length
*      18052      18151: gap of unknown length
*      18152      19373: contig of 1222 bp in length
*      19374      19473: gap of unknown length
*      19474      20668: contig of 1195 bp in length
*      20669      20768: gap of unknown length
*      20769      22399: contig of 1631 bp in length
*      22400      22499: gap of unknown length
*      22500      24200: contig of 1701 bp in length
*      24201      24300: gap of unknown length
*      24301      25667: contig of 1367 bp in length
*      25668      25767: gap of unknown length
*      25768      27705: contig of 1938 bp in length
*      27706      27805: gap of unknown length

```

FEATURES

source

```

*      27806      29643: contig of 1838 bp in length
*      29644      29743: gap of unknown length
*      29744      31718: contig of 1975 bp in length
*      31719      31818: gap of unknown length
*      31819      33299: contig of 1481 bp in length
*      33300      33399: gap of unknown length
*      33400      35588: contig of 2189 bp in length
*      35589      35688: gap of unknown length
*      35689      37969: contig of 2280 bp in length
*      37969      38068: gap of unknown length
*      38069      40341: contig of 2273 bp in length
*      40342      40441: gap of unknown length
*      40442      42574: contig of 2133 bp in length
*      42575      42674: gap of unknown length
*      42675      44938: contig of 2264 bp in length
*      44939      45038: gap of unknown length
*      45039      50126: contig of 5088 bp in length
*      50127      50226: gap of unknown length
*      50227      52594: contig of 2368 bp in length
*      52595      52695: gap of unknown length
*      52696      57156: contig of 4462 bp in length
*      57157      57256: gap of unknown length
*      57257      61310: contig of 4054 bp in length
*      61311      61410: gap of unknown length
*      61411      65807: contig of 4397 bp in length
*      65808      65908: gap of unknown length
*      65909      70160: contig of 4253 bp in length
*      70161      70260: gap of unknown length
*      70261      76920: contig of 6659 bp in length
*      76920      77020: gap of unknown length
*      77020      85765: contig of 8745 bp in length
*      85766      85865: gap of unknown length
*      85866      106630: contig of 20766 bp in length
*      106631      106730: gap of unknown length
*      106731      134020: contig of 27290 bp in length
*      134021      134120: gap of unknown length
*      134121      165082: contig of 30962 bp in length
*      165083      165182: gap of unknown length
*      165183      196240: contig of 31055 bp in length
*      196241      196340: gap of unknown length
*      196341      232180: contig of 35840 bp in length.
*      232180      Location/Qualifiers
1. .232180
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="1"
  /clone="RP11-588H15"
1. .1312
  /note="assembly_name:Contig14"
1413. .2446
  /note="assembly_name:Contig15"
2547. .3878
  /note="assembly_name:Contig30"
3979. .5215
  /note="assembly_name:Contig35"
5316. .6544
  /note="assembly_name:Contig40"
6645. .7940
  /note="assembly_name:Contig42"
8041. .9190
  /note="assembly_name:Contig44"
9291. .10909
  /note="assembly_name:Contig47"
11010. .12203
  /note="assembly_name:Contig48"
12304. .13542
  /note="assembly_name:Contig52"
13643. .14982
  /note="assembly_name:Contig53"
15083. .16905
  /note="assembly_name:Contig54"
17006. .18051
  /note="assembly_name:Contig55"

```


TITLE Gastrointestinal defensins, cDNA sequences and method for the
JOURNAL Production and use thereof
Patent: US 5641497-A 6/24/JUN-1997,
FEATURES Location/Qualifiers
source 1..452
/organism="unknown"
BASE COUNT 124 a 118 c 98 g 112 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0213 Length: 452
Score: 64.00 Matches: 15
Percent Similarity: 89.47% Conservative: 2
Best Local Similarity: 78.95% Mismatches: 2
Query Match: 71.11% Indels: 0
Gaps: 0
US-10-045-180A-4 (1-19) x I49588 (1-452)
Qy 1 MetAgtTtTleuThrlleuSerAlaPheleuLeuValAlaLeuGlnAlaTtTpAla 19
Db 19 ATGAGAACCCCTCACCACCTCCTGCTCTCTCTCGGCGCCCTCCAGGCCAAGGCT 75
RESULT 10
LOCUS HUMDFSNSIX 452 bp mRNA linear PRI 27-APR-1993
DEFINITION Homo sapiens defensin 6 mRNA, complete cds.
ACCESSION M98331
VERSION M98331.1 GI:181546
KEYWORDS defensin 6
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 452)
Jones, P.B. and Bevins, C.L.
Defensin-6 mRNA in human Paneth cells: implications for
antimicrobial peptides in host defense of the human bowel
JOURNAL FEBS Lett. 315 (2), 187-192 (1993)
MEDLINE 93114459
PUBMED 841977
COMMENT Original source text: Homo sapiens small intestine cDNA to mRNA.
FEATURES Location/Qualifiers
source 1..452
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="small intestine"
5'UTR 1..18
CDS 19..321
/note="prepropeptide"
/codon_start=1
/product="defensin 6"
/protein_id="AAB59357.1"
/db_xref="GI:181547"
/translation="MRTITLITAVLLVALQAKAEPLQAEEDPLQAKAYEADAQEOGGA
NDOPFAVSFAEDASSSLRALGSTRAPFCHCRRCSTREYSGTCTWGINHRCCL"
3'UTR 319..440
variation 390
/replace="t"
BASE COUNT 124 a 118 c 98 g 112 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0213 Length: 452
Score: 64.00 Matches: 15
Percent Similarity: 89.47% Conservative: 2
Best Local Similarity: 78.95% Mismatches: 2
Query Match: 71.11% Indels: 0
Gaps: 0
US-10-045-180A-4 (1-19) x HUMDFSNSIX (1-452)

Qy 1 MetAgtTtTleuThrlleuSerAlaPheleuLeuValAlaLeuGlnAlaTtTpAla 19
Db 19 ATGAGAACCCCTCACCACCTCCTGCTCTCTCTCGGCGCCCTCCAGGCCAAGGCT 75
RESULT 11
LOCUS HSU33317 3060 bp DNA linear PRI 23-FEB-1996
DEFINITION Human defensin 6 (HD-6) gene, complete cds.
ACCESSION U33317
VERSION U33317.1 GI:1200181
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 3060)
Mallow, E.B., Harris, A., Salzman, N., Russell, J.P.,
Debernardinis, R.U., Kuchell, E. and Bevins, C.L.
Human enteric defensins. Gene structure and developmental
expression
JOURNAL J. Biol. Chem. 271 (8), 4038-4045 (1996)
MEDLINE 96223969
PUBMED 8626737
REFERENCE 2 (bases 1 to 3060)
Charles L. Bevins.
Direct Submission
Submitted (07-AUG-1995) Charles L. Bevins, Pediatrics, The
Children's Hospital of Philadelphia, 34th and Civic Center Blvd.,
Philadelphia, PA 19104, USA
FEATURES Location/Qualifiers
source 1..3060
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
CHAT_signal 1278
TATA_signal 1341
gene 1370..2730
/gene="HD-6"
/join(1370..1605,2517..2730)
mRNA 1370..1410
/gene="HD-6"
5'UTR 1370..1410
/gene="HD-6"
/join(1411..1605,2517..2627)
CDS 1370..1410
/gene="HD-6"
/codon_start=1
/product="defensin 6"
/protein_id="AAC30382.1"
/db_xref="GI:1200182"
/translation="MRTITLITAVLLVALQAKAEPLQAEEDPLQAKAYEADAQEOGGA
NDOPFAVSFAEDASSSLRALGSTRAPFCHCRRCSTREYSGTCTWGINHRCCL"
3'UTR 2628..2730
/gene="HD-6"
polyA_signal 2730
/gene="HD-6"
BASE COUNT 892 a 678 c 612 g 878 t
ORIGIN
Alignment Scores:
Pred. No.: 0.158 Length: 3060
Score: 64.00 Matches: 15
Percent Similarity: 89.47% Conservative: 2
Best Local Similarity: 78.95% Mismatches: 2
Query Match: 71.11% Indels: 0
Gaps: 0
US-10-045-180A-4 (1-19) x HSU33317 (1-3060)
Qy 1 MetAgtTtTleuThrlleuSerAlaPheleuLeuValAlaLeuGlnAlaTtTpAla 19
Db 1411 ATGAGAACCCCTCACCACCTCCTGCTCTCTCTCGGCGCCCTCCAGGCCAAGGCT 1467
RESULT 12

AF314060 79730 bp DNA linear PRI 09-JUL-2002
LOCUS Homo sapiens chromosome 8 clone Qu-L-12046 map 8p22-p21, complete
DEFINITION
ACCESSION AF314060
VERSION AF314060
KEYWORDS
SOURCE HTG. GI:21717333
ORGANISM Homo sapiens (human)
REFERENCE
AUTHORS
TITLE Chromosome 8 genomic sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 79730)
AUTHORS Genome Sequencing Center Jena.
TITLE Direct Submissions
JOURNAL Submitted (17-OCT-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
3 (bases 1 to 79730)
AUTHORS Lagemann,D. and Platzer,M.
TITLE Direct Submissions
JOURNAL Submitted (09-JUL-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On Jul 9, 2002 this sequence version replaced gi:1098844.
COMMENT
----- Genome Center
Center: Institute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gscj-submit@genome.imb-jena.de
----- Project Information
Center project name: H376
Center clone name: Qu-L-12046
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 79710 bases at least Q40
Consensus quality: 79727 bases at least Q30
Consensus quality: 79729 bases at least Q20
Quality coverage: 12.19x

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one pUC18 subclone; and the
assembly was confirmed by restriction digest.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES
source
1. 79730
location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p22-p21, complete sequence."
/clone="Qu-L-12046"
172
variation

/note="C substituted in clone: GSI-P1-861"
/replace="A"
500
/note="T substituted in clone: SCB-633e22, Qu-L-12046"
/replace="C"
679
/note="G substituted in clone: Qu-L-12046, SCB-633e22"
/replace="A"
729
/note="T substituted in clone: Qu-L-12046, SCB-633e22"
/replace="G"
791
/note="A substituted in clone: Qu-L-12046, GSI-P1-861,
SCB-633e22"
/replace="G"
1245
/note="C substituted in clone: SCB-633e22, GSI-P1-861,
Qu-L-12046"
/replace="T"
1668
/note="A substituted in clone: Qu-L-12046, SCB-633e22,
GSI-P1-861"
/replace="G"
1777
/note="G substituted in clone: GSI-P1-861"
/replace="T"
1866
/note="A substituted in clone: Qu-L-12046, SCB-633e22"
/replace="G"
2621
/note="C substituted in clone: GSI-P1-861"
/replace="T"
3021..3022
/note="deleted in clone: SCB-633e22, GSI-P1-861"
/replace=""
3284
/note="G substituted in clone: SCB-633e22"
/replace="A"
3894
/note="A substituted in clone: SCB-633e22"
/replace="T"
4041
/note="A substituted in clone: GSI-P1-861"
/replace="G"
4204
/note="deleted in clone: GSI-P1-861, SCB-633e22,
CMD-2629116"
/replace=""
4510
/note="T substituted in clone: Qu-L-12046, SCB-633e22"
/replace="C"
4808
/note="C substituted in clone: Qu-L-12046, SCB-633e22"
/replace="T"
4949
/note="A substituted in clone: Qu-L-12046, SCB-633e22"
/replace="G"
5121
/note="T substituted in clone: Qu-L-12046, SCB-633e22"
/replace="C"
5142
/note="A substituted in clone: GSI-P1-861"
/replace="G"
6261
/note="T substituted in clone: Qu-L-12046, SCB-633e22"
/replace="C"
6473
/note="T substituted in clone: SCB-633e22"
/replace="G"
6870
/note="C substituted in clone: GSI-P1-861"
/replace="T"
7648
variation

All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

FEATURES

Location/Qualifiers

```

1.150940
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="p22-p21"
/clone="CTD-2629116"
31743..32232
/note="single stranded/single chemistry region"
misc_feature
38626..38692
/note="single stranded/single chemistry region"
misc_feature
93307..93310
/note="low quality region"
misc_feature
93393..93395
/note="low quality region"
misc_feature
93427..93428
/note="low quality region"
misc_feature
93434..93436
/note="low quality region"
misc_feature
93435
/note="low quality region"
misc_feature
93481
/note="low quality region"
misc_feature
93481
/note="low quality region"
misc_feature
93534
/note="low quality region"
misc_feature
93534
/note="low quality region"
BASE COUNT 39368 a 33223 c 36391 g 41958 t
ORIGIN

```

Alignment Scores:

```

Pred. No.: 9.47 Length: 150940
Score: 64.00 Matches: 15
Percent Similarity: 89.47% Conservative: 2
Best Local Similarity: 78.95% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 9 Gaps: 0

```

US-10-045-180A-4 (1-19) X AF233439 (1-150940)

QY 1 MetAgtThreutThreutUserAlaphleuleuValAlaleuglnAlatpAla 19

Db 85970 ATGAGAACCTCACAACATCTGCTTCTCTCTGCGCCCTCCAGCCAGGCT 85914

RESULT 14 AF200455/c 158653 bp DNA linear PRI 09-JAN-2003

LOCUS Homo sapiens chromosome 8 clone SCB-540n10 map p22-p21, complete

DEFINITION sequence.

ACCESSION AF200455

VERSION AF200455.3 GI:27552755

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 158653)

AUTHORS Schudy,A., Blechschmidt,K., Menzel,U., Polley,A., Reichwald,K., Rump,A., Schlhabel,M.B., Taudien,S., Wen,G., Schutte,B., Malik,M., Peng,J., Hong,J., McCray,J., and Rosenthal,A.

TITLE Chromosome 8 genomic sequence

REFERENCE 2 (bases 1 to 158653)

AUTHORS Schudy,A., Schlhabel,M., Schutte,B., Malik,M., Hong Peng,J., McCray,J., Baumgart,C., Menzel,U., Weber,J., Schattevoy,R. and Rosenthal,A.

TITLE Direct Submission

JOURNAL Submitted (28-Oct-1999) Genome Analysis, Institute of Molecular

REFERENCE Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

AUTHORS 3 (bases 1 to 158653)

TITLE Genome Sequencing Center Jena.

JOURNAL Direct Submission

Submitted (24-May-2000) Genome Analysis, Institute of Molecular

REFERENCE Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

AUTHORS 4 (bases 1 to 158653)

TITLE lagemann,D. and Platzner,M.

JOURNAL Direct Submission

Submitted (09-JAN-2003) Genome Analysis, Institute of Molecular

REFERENCE Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

AUTHORS On Jan 9, 2003 this sequence version replaced gi:8151988.

TITLE Genome Center

Center: Institute of Molecular Biotechnology

Center code: IMB

Web site: http://genome.imb-jena.de/

Contact: gscj-submit@genome.imb-jena.de

Project Information

Center project name: H293

Center clone name: SCB-540n10

Summary Statistics

Sequencing vector: pUC18; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 157265 bases at least Q40

Consensus quality: 158014 bases at least Q30

Consensus quality: 158652 bases at least Q20

Quality coverage: 8.15x

This sequence was finished as follows unless otherwise noted: all

regions were double stranded, sequenced with an alternate

chemistry, or covered by high quality data (i.e., phred quality >=

30): an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the

assembly was confirmed by restriction digest.

Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

FEATURES

source

Location/Qualifiers

```

1.158653
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="p22-p21"
/clone="SCB-540n10"
773..805
/note="single stranded/single chemistry region"
misc_feature
1550..1558
/note="low quality region"
misc_feature
1554
/note="single stranded/single chemistry region"
misc_feature
1557..1596
/note="single stranded/single chemistry region"
misc_feature
1571
/note="low quality region"
misc_feature
2687..3061
/note="single stranded/single chemistry region"
misc_feature
14120..14203
/note="single stranded/single chemistry region"
misc_feature
25023
/note="single stranded/single chemistry region"
misc_feature
25552
misc_feature
25774

```

```

misc_feature /note="low quality region"
26465..26484
/note="single stranded/single chemistry region"
misc_feature 26485..26616
/note="single clone coverage"
misc_feature 26592
/note="low quality region"
unsure 32723
unsure 38261
unsure 42450..42453
unsure 43111..43237
misc_feature /note="single stranded/single chemistry region"
43134
unsure 43160
unsure 43183
unsure 43191
misc_feature 46094..46149
/note="single stranded/single chemistry region"
unsure 46189
unsure 46238..46240
misc_feature 48234..48285
/note="single stranded/single chemistry region"
misc_feature 53580..53612
/note="single stranded/single chemistry region"
misc_feature 53587
/note="low quality region"
misc_feature 53589
/note="low quality region"
misc_feature 53595..53597
/note="low quality region"
unsure 53595..53597
misc_feature 53613..53709
/note="single clone coverage"
misc_feature 53641..53645
/note="low quality region"
misc_feature 53649..53655
/note="low quality region"
unsure 53651
misc_feature 53660..53661
/note="low quality region"
misc_feature 53666..53675
/note="low quality region"
unsure 53672
unsure 53688
misc_feature 53696..53709
/note="low quality region"
unsure 53699
misc_feature 63788
/note="low quality region"
misc_feature 63791
/note="low quality region"
misc_feature 63816..63822
/note="low quality region"
misc_feature 63836
/note="low quality region"
misc_feature 63842
/note="low quality region"
misc_feature 63854
/note="low quality region"
misc_feature 63860
/note="low quality region"
unsure 63860
misc_feature 63873..63877
/note="low quality region"
misc_feature 63882..63889
/note="low quality region"
misc_feature 63898..63903
/note="low quality region"
misc_feature 63916..63917
/note="low quality region"
misc_feature 63927..64074
/note="single stranded/single chemistry region"
misc_feature 63929..63985

```

```

unsure /note="low quality region"
63958
unsure 63961
unsure 63982
misc_feature 64015..64017
/note="low quality region"
misc_feature 64030..64036
/note="low quality region"
misc_feature 64045..64048
/note="low quality region"
misc_feature 73345..73421
/note="single stranded/single chemistry region"
misc_feature 75927..76057
/note="single stranded/single chemistry region"
misc_feature 83076..83120
/note="single stranded/single chemistry region"
misc_feature 83099..83182
/note="low quality region"
misc_feature 83121..83206
/note="single clone coverage"
unsure 83123
unsure 83143
unsure 83147
misc_feature 83207..83309
/note="single stranded/single chemistry region"
misc_feature 83243..83245
/note="low quality region"
misc_feature 84552
/note="low quality region"
misc_feature 84556..84700
/note="single stranded/single chemistry region"
misc_feature 84589..84590
/note="low quality region"
misc_feature 84607..84610

Alignment Scores:
Pred. No.: 9.97
Score: 64.00
Percent Similarity: 89.47%
Best Local Similarity: 78.95%
Query Match: 71.11%
DB: 9 Gaps: 0

US-10-045-180A-4 (1-19) X AF200455 (1-158653)

Cy 1 MetArchthireuthireulenseRAlaphleuleuValAlaengInAlATpAla 19
Db 29175 ATGAGAACCTCCACACCTCCTGCTCTCTCTGCGCCCTCCAGGCCAGGCT 29119

RESULT 15
AC079018/c 168702 bp DNA linear PRI 15-AUG-2002
LOCUS Homo sapiens chromosome 8, clone RP11-161B1, complete sequence.
DEFINITION AC079018
ACCESSION AC079018.10 GI:22263526
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 168702)
Homo sapiens chromosome 8, clone RP11-161B1
Unpublished
2 (bases 1 to 168702)
REFERENCES
AUTHORS Anderson,S., Barna,N., Bastien,V., Bede,F., Bogunlavkiy,L.,
Birren,B., Linton,L., Nusbaum,C., Lander,B., Abraham,H., Allen,N.,
Bookbinder,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
DeRellano,K., Dewar,K., Diaz,U.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRocque,K.,

```

Lamazares, R., Landers, T., Lehocaky, J., Levine, R., Lieu, C., Liu, G., MacDonald, P., Margulis, N., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivari, T.M., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Rieback, M., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Soudneaz, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Sudramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigglio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS
 Direct Submission
 Submitted (15-AUG-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 168702)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, V., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Plunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS
 Direct Submission
 Submitted (06-ANG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 168702)

Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Plunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
 Direct Submission
 Submitted (15-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 15, 2002 this sequence version replaced gi:22123231.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WITBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L10753
 Center clone name: 161_B_1
 ----- Location/Qualifiers

source	
1. 168702	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="8"
	/map="8"
	/clone="RP11-161B1"
	/clone_lib="RP11-11 Human Male BAC"
unsure	
repeat_region	/note="PCR product sequence only"
repeat_region	/note="471
repeat_region	/rpt_family="(TG)n"
repeat_region	528..631
repeat_region	/rpt_family="(TG)n"
unsure	832..836
repeat_region	/note="<30 qual SNGL region"
repeat_region	2098..2371
repeat_region	/rpt_family="AluJo"
repeat_region	2909..3311
repeat_region	/rpt_family="L2"
repeat_region	3394..3422
repeat_region	/rpt_family="AT rich"
repeat_region	complement(3799..4104)
repeat_region	/rpt_family="AluSx"
repeat_region	complement(4115..4425)
repeat_region	/rpt_family="AluY"
repeat_region	4907..4934
repeat_region	/rpt_family="(TTTTTC)n"
repeat_region	6023..6326
repeat_region	/rpt_family="AluY"
repeat_region	7266..7764
repeat_region	/rpt_family="MER74B"
repeat_region	7767..8016
repeat_region	/rpt_family="AluSx"
repeat_region	10421..10478
repeat_region	/rpt_family="(TA)n"
repeat_region	complement(10482..10962)
repeat_region	/rpt_family="MER31B"
repeat_region	complement(10984..11185)
repeat_region	/rpt_family="MS7A"
repeat_region	11218..11522
repeat_region	/rpt_family="AluSx"
repeat_region	11733..11851
repeat_region	/rpt_family="L1M4"
repeat_region	complement(11919..12193)
repeat_region	/rpt_family="AluSx"
repeat_region	12434..12735
repeat_region	/rpt_family="AluY"
repeat_region	12761..13748
repeat_region	/rpt_family="L1M4"
repeat_region	complement(13875..14169)
repeat_region	/rpt_family="AluSx"
repeat_region	14995..15035
repeat_region	/rpt_family="(TGA)n"
repeat_region	15872..16053
repeat_region	/rpt_family="L2"
repeat_region	17044..17211
repeat_region	/rpt_family="MLT1H"
repeat_region	17301..17598
repeat_region	/rpt_family="MLT1B"
repeat_region	complement(18497..18646)
repeat_region	/rpt_family="MIR"
repeat_region	complement(19356..24412)
repeat_region	/rpt_family="L1P45"
repeat_region	complement(24410..24506)
repeat_region	/rpt_family="L1P45"
repeat_region	24389..25170
repeat_region	/rpt_family="(CA)n"
repeat_region	25202..25467
repeat_region	/rpt_family="(CA)n"
repeat_region	25486..25588
repeat_region	/rpt_family="(CA)n"
repeat_region	25613..25784

FEATURES

```
repeat_region      /rpt_family="(CA)n"
                    25831..25969
repeat_region      /rpt_family="(CAT)n"
                    25949..26079
repeat_region      /rpt_family="(CA)n"
                    26481..26856
repeat_region      /rpt_family="MER115"
                    complement(27532..27905)
repeat_region      /rpt_family="MER4A"
                    complement(27906..28004)
repeat_region      /rpt_family="MER4A"
                    complement(28176..28667)
repeat_region      /rpt_family="MER74A"
                    29190..29439
repeat_region      /rpt_family="L2"
                    complement(29440..29898)
repeat_region      /rpt_family="MLT1C"
                    29899..29971
repeat_region      /rpt_family="L2"
                    30095..30125
repeat_region      /rpt_family="(GAATG)n"
                    30647..30668
repeat_region      /rpt_family="AT_rich"
                    complement(31093..31397)
repeat_region      /rpt_family="AluSp"
                    31444..31570
repeat_region      /rpt_family="L1M4"
                    31821..32123
repeat_region      /rpt_family="AluSx"
                    32692..32984
repeat_region      /rpt_family="AluSc"
                    complement(33681..33913)
```

```
Alignment Scores:
Pred. No.:          10.6          Length:      168702
Score:             64.00          Matches:      15
Percent Similarity: 89.47%          Conservative: 2
Best Local Similarity: 78.95%          Mismatches: 2
Query Match:       71.11%          Indels:      0
DB:                9              Gaps:          0
```

US-10-045-180A-4 (1-19) x AC079018 (1-168702)

```
Qy      1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
          |||||:::|||||
Db      100083 ATGAGAACCTCACCATCTCACTGCTCTCCTGTCGCCCTCAGGCCAAGGCT 100027
```

Search completed: December 17, 2003, 16:37:39
Job time : 603.622 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 664.293 Seconds
(without alignments)
695.153 Million cell updates/sec

Title: US-10-045-180a-4
Perfect score: 90
Sequence: 1 MRTLTLASFLVALQAWA 19

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=/cgn2.1/USPTO.spool.p/US10045180/runat.17122003.145024.12540/arp_query.fasta_1.860
-DB=EST -QFMT=fastap -SUFFIX=rcs -MINMATCH=0.1 -IOFCL=0 -IOOEXT=0 -LIST=45
-UNITS=bits -START=1 -END=1 -MATRIX=bitsum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR SCORE=DCT -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10045180 @cgn 1.1 6100 @runat.17122003.145024.12540 -NCPU=6 -ICPU=3
-NO MAP -LARGEBUFFER -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAP TIMEOUT=30 -THREADS=1 XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:
1: em_estba:*
2: em_esthum:*
3: em_estcin:*
4: em_estcnu:*
5: em_estcov:*
6: em_estpl:*
7: em_estro:*
8: em_hrc:*
9: gb_estc1:*
10: gb_estc2:*
11: gb_hrc:*
12: gb_estc3:*
13: gb_estc4:*
14: gb_estc5:*
15: em_estcun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrt1:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	70.0	221	10	BF895944 CM2-MT015
2	61	67.8	168	12	BF895944 CM2-MT015
3	61	67.8	194	10	BF895944 CM2-MT015
4	61	67.8	243	13	BF895944 CM2-MT015
5	61	67.8	251	9	BF895944 CM2-MT015
6	61	67.8	281	9	BF895944 CM2-MT015
7	61	67.8	289	9	BF895944 CM2-MT015
8	61	67.8	290	10	BF895944 CM2-MT015
9	61	67.8	294	9	BF895944 CM2-MT015
10	61	67.8	298	9	BF895944 CM2-MT015
11	61	67.8	312	9	BF895944 CM2-MT015
12	61	67.8	320	9	BF895944 CM2-MT015
13	61	67.8	327	9	BF895944 CM2-MT015
14	61	67.8	348	10	BF895944 CM2-MT015
15	61	67.8	349	12	BF895944 CM2-MT015
16	61	67.8	349	12	BF895944 CM2-MT015
17	61	67.8	353	10	BF895944 CM2-MT015
18	61	67.8	354	9	BF895944 CM2-MT015
19	61	67.8	354	9	BF895944 CM2-MT015
20	61	67.8	355	10	BF895944 CM2-MT015
21	61	67.8	360	12	BF895944 CM2-MT015
22	61	67.8	370	10	BF895944 CM2-MT015
23	61	67.8	372	10	BF895944 CM2-MT015
24	61	67.8	377	9	BF895944 CM2-MT015
25	61	67.8	379	10	BF895944 CM2-MT015
26	61	67.8	379	12	BF895944 CM2-MT015
27	61	67.8	387	12	BF895944 CM2-MT015
28	61	67.8	388	10	BF895944 CM2-MT015
29	61	67.8	389	12	BF895944 CM2-MT015
30	61	67.8	396	10	BF895944 CM2-MT015
31	61	67.8	413	9	BF895944 CM2-MT015
32	61	67.8	413	9	BF895944 CM2-MT015
33	61	67.8	424	10	BF895944 CM2-MT015
34	61	67.8	425	10	BF895944 CM2-MT015
35	61	67.8	430	9	BF895944 CM2-MT015
36	61	67.8	430	9	BF895944 CM2-MT015
37	61	67.8	430	9	BF895944 CM2-MT015
38	61	67.8	433	10	BF895944 CM2-MT015
39	61	67.8	449	10	BF895944 CM2-MT015
40	61	67.8	450	10	BF895944 CM2-MT015
41	61	67.8	465	9	BF895944 CM2-MT015
42	61	67.8	472	10	BF895944 CM2-MT015
43	61	67.8	473	10	BF895944 CM2-MT015
44	61	67.8	475	10	BF895944 CM2-MT015
45	61	67.8	486	9	BF895944 CM2-MT015

ALIGNMENTS

RESULT 1
BF895944/c
LOCUS CM2-MT015-221100-548-b02 MT0157 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF895944
VERSION BF895944.1 GI:12287403
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 221)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

TITLE	Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
PUBMED	10737800
COMMENT	Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpcson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?rl=CM2&t2=CM2-MT0157-221100-548-b02&t3=2000-11-22&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 54
High quality sequence stop: 221.

FEATURES	SOURCE	Location/Qualifiers
1.	.221	
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/dev_stage="Adult"
		/clone_id="MT0157"
		/note="Organ: marrow; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. letters Patent application No. 197,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT	43 a	64 c 68 g 46 t
ORIGIN		

Alignment Scores:	
Pred. No.:	0.849
Score:	63.00
Percent Similarity:	84.21%
Best Local Similarity:	78.95%
Query Match:	70.00%
DB:	10
US-10-045-180A-4 (1-19)	x BF895944 (1-221)
	Length: 221
	Matches: 15
	Conservative: 1
	Mismatches: 3
	Indels: 0
	Gaps: 0

Qy	Db	RESULT 2	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS
1	140	ATGAGGACCCCTGCACCTCTTCTGTCGCCATTCTCCTGATGGCCCTGCAGAGCCCAAGCT	BI023130	168 bp	CM0-MT0390-070201-767-b09	MT0390	Homo sapiens	CDNA, mRNA	linear	EST 14-JUN-2001	
					BI023130						
					BI023130.1		GI:14429760				
					EST.						
					Homo sapiens (human)						
					Homo sapiens						
					Bukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;						
					Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.						
					1 (bases 1 to 168)						
					Dias Neto,E., Garcia Correa,R., Vertovskii-Almeida,S., Birones,M.R.,						

Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsumura, A., Bata, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and

TITLE	Simpson, A.J. Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3451-3496 (2000)
PUBMED	20202663
COMMENT	Contact: Simpson A.J.G. 10737800

Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel : +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL:
<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM0&t2=CM0-MT0390>
070201-767-b09&t3=2001-02-07&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 15
High quality sequence stop: 168.
CGAGGATGGC

```

1. 1.168
SOURCE
Location/Qualifiers
1. 1.168
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_id="WT0390"
/note="Organ: marrow; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 1997, 716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT
30 a 59 c 43 g 36 t

```

```

Alignment Scores:
Pred. No.:      1.21      Length:      168
Score:          61.00      Matches:     14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches:    3
Query Match:    67.78%   Indels:         0
DB:             12       Gaps:          0

```

QY	1	MesArgThrIreuthrIreuleuSerAlaPheLeuIleValAlaLeuGlnAlaTrpAla	19
		:::::	
Dd	101	ATAGAGACCTCGCCCATCCTTGTCGCATTCTCTGTGGCCCTGCAGGCCAGGCT	157
RESULT 3			
BGI93795			
LOCUS	BGI93795	194 bp	mRNA linear EST 21-APR-2002
DEFINITION	RST119311	Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.	
ACCSSION	BGI93795		
VERSION	BGI93795.1	GI:13715482	

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Karyoyuca; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (Phases 1 to 194)
AUTHORS	Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R., Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J., Lerner, L., Costanzo, D., Melligott, K., Boozar, S., Mays, R., Smith, E., Veloso, N., Klitz, A., Hess, J., Cochren, K., Lo, K., Offenbacher, J., Danzig, J. and Ducar, M.
TITLE	Creation of genome-wide protein expression libraries using random activation of gene expression
JOURNAL	Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE	21227151

Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research

The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org

For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/cdb/hgi/hgi.html>)
 Seq primer: M13 Reverse..

```

FEATURES
source
location/Qualifiers
1..281
/organism="Homo sapiens"
/mol_type="mRNA"
/od_xref="ATCC (inhost):121891"
/db_xref="taxon:9606"
/sex="mixed"
/tissue_type="bone marrow"
/dev_stage="adult"
/clone_id="Bone marrow"
/note="Vector: pBluescript SK-

```

BASE COUNT	60 a	87 c	78 g	56 t
ORIGIN				
Alignment Scores:				
Pred. No.:	2.52		Length:	281
Score:	61.00		Matches:	14
Percent Similarity:	84.21%		Conservative:	2

Query Match:	67.78%
2000 Local Similarity:	73.00%
Mil	Inq

US-10-045-180A-4 (1-19) x AA321281 (1-281)

06210589CC11C6CACA1C11C1C1C1G1G3CCCGCAGAGCCCAAGCT 124

LOCUS	AA321317	289 bp	mRNA	linear	EST 19-APR-1997
DEFINITION	ESR232184 Bone marrow				
ACCESSION	AA321317				
VERSION	AA321317.1	GI:1973644			
KEYWORDS	EST.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				

REFERENCE

AUTHORS

1 (pages 1 to 289)

Adams, W.D., Kertlaige, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-ai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geobagasan, N.S., Glodde, A., Ghehm, C.L., Hanna, M.C., Hedblom, E., Hinkley, P.S. Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmoreo, S.M., Merrick, J.M., Moreno-Beltranque, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Sauder, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utechtack, T.R., Weidman, J.P., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, M.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Konak, D.L., Kunsch, C., Hungjun, J., Li, H., Weissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Whang, J., Xu, C., Yu, G.L., Ruben, S.M., Dillon, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M. and Venter, J.C.

TITLE

Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

Nature 377 (6542 Suppl) 3:374 (1995)

Adams, M.D., Verklavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kitchens, E.F., Weinstein, K.G., Gossayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle, H.J., Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fitchman, J.L., Geogaasen, N.S., Glosck, A., Gnahm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S. Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Metrick, J.M., Moreno-Palanges, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, K.A., Ryder, S.B., Scott, J.L., Sauder, D.M., Shirley, R., Small, K.V., Springs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.U., Dinke, D., Feng, D.-F., Ferrite, A., Fischer, C., Hastings, G.A., He, W.W., Hu, T.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L.,

Eukaryota, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 327)
Adams, M.D., Kewalavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bolt, C.J., Lee, N.H., Kirnesh, E.F., Weinsmann, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Weil, C., Claydon, R.A., Lin, T.R., Cotton, M.D., Easte-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitchugh, W.M., Fitchman, J.L., Geoghegan, N.S., Glodok, A., Guelm, C.L., Hanna, M.C., Hedlow, E., Hinkle, P.S.Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaro, S.M., Merrick, J.M., Moreno-Palancques, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, T.L., Saudak, D.M., Shirley, R., Small, K.V., Springs, T.A., Utechtack, T.R., Weidman, J.F., Li, Y.

TITLE
Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence

JOURNAL MEDLINE PUBMED
Nature 377 (6547 Suppl), 3-174 (1995)
756098

COMMENT
Other_BESTs: THC169174
Contact: Kerlavage, AR
Bioinformatics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850 USA
Tel: 3018699056
Fax: 3018699423
Email: arkerlav@tigr.org
For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
Seq primer: M13 Reverse.

FEATURES
source
Location/Qualifiers

1..327
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):122063"
/db_xref="taxon:9606"
/sex="mixed"
/tissue_type="bone marrow"
/dev_stage="adult"
/clone_id="Bone marrow"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"

BASE COUNT
ORIGIN 72 a 97 c 91 g 66 t 1 others

Alignment Scores:
Pred. No.: 3.13 Length: 327
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 9 Gaps: 0

US-10-045-180a-4 (1-19) x AA321192 (1-327)

Qy 1 MetAgtThrLeuThrIleuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
|||||:::|||||
Db 75 ATGAGGACCCCTGCCATCTTCCTGCTGCGCCCTGCAGGCCAGGCT 131

RESULT 14
BF893385 348 bp mRNA linear EST 18-JUN-2001
LOCUS QV3-MT0129-111100-427-h01 MT0129 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF893385
VERSION BF893385.1 GI:12284484
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 348)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.G.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL MEDLINE PUBMED
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV3&2=QV3-MT0129-111100-427-h01&3=2000-11-11&4=1>)
Seq primer: puc 18 forward
High quality sequence stop: 348.

FEATURES
source
Location/Qualifiers

1..348
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_id="MT0129"
/note="Organ: marrow; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
ORIGIN 79 a 103 c 97 g 69 t

Alignment Scores:
Pred. No.: 3.42 Length: 348
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 10 Gaps: 0

US-10-045-180a-4 (1-19) x BF893385 (1-348)

Qy 1 MetAgtThrLeuThrIleuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
|||||:::|||||
Db 87 ATGAGGACCCCTGCCATCTTCCTGCTGCGCCCTGCAGGCCAGGCT 143

RESULT 15
BI021617 349 bp mRNA linear EST 14-JUN-2001
LOCUS BI021617/c
DEFINITION CM3-MT0343-170101-665-805 MT0343 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI021617
VERSION BI021617.1 GI:14428247
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
1 (bases 1 to 349)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.G.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL MEDLINE PUBMED
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 71.25 Seconds
(without alignments)
719.850 Million cell updates/sec

Title: US-10-045-180A-4
Sequence: 1 MRTLITLSAFLVALQAMA 19

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xip
-O=cpm2.1/USPRO.spool.p/US10045180/runat.17122003.145023.12520/app.query.fasta.1.860
-DB=N Geneseg.19Jun03 -QPM=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPT=0
-LIST=45 -DOCMATCH=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human0.cdi
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USR=US10045180 @CGN.1.1.835 @runat.17122003.145023.12520 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseg.19Jun03:*

1: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseg/geneseg-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	90	100.0	453	20	AAK26697	CDNA sequence enco
2	90	100.0	453	20	AAK26698	CDNA sequence enco
3	90	100.0	415	20	AAK26696	Nucleotide sequenc
4	64	71.1	294	22	AAH57220	Human intestine sp
5	64	71.1	398	24	ABN59997	Novel human coding
6	64	71.1	452	14	AAO53217	Sequence encoding
7	64	71.1	452	25	ACA56600	Human signalling p
8	61	67.8	110	17	AAK33496	110mer used in con
9	61	67.8	110	18	AAK63059	Human pre-pro cart
10	61	67.8	110	19	AAV10895	Human pre-pro cart
11	61	67.8	187	22	ABA75757	Human foetal liver
12	61	67.8	187	22	AAK50398	Human bone marrow
13	61	67.8	187	24	AB523860	Human genome-deriv
14	61	67.8	229	16	AAK26106	Human gene signatu
15	61	67.8	464	21	AAK20911	Human defensin 3 p
16	61	67.8	464	21	AAK34789	Human adenosine re
17	61	67.8	478	21	AAK04468	Human secreted pro
18	61	67.8	498	12	AAQ14808	Chronic myelogenou
19	61	67.8	498	24	AB080601	Human defensin 1 c
20	61	67.8	514	24	ABN59722	Novel human coding
21	61	67.8	597	22	ABA63315	Human foetal liver
22	61	67.8	597	22	AAK37526	Human bone marrow
23	61	67.8	597	24	AB511502	Human genome-deriv
24	61	67.8	644	21	AAK21446	Human defensin 3 p
25	61	67.8	652	24	AB211202	Human polynucleoti
26	61	67.8	664	24	AB211183	Human polynucleoti
27	61	67.8	3710	21	AAK20910	Human defensin 3 p
28	61	67.8	3710	21	AAK34788	Human adenosine re
29	61	67.8	3710	24	AB080602	Human defensin 1-r
30	61	67.8	3710	24	ABK83829	Human CDNA differe
31	61	67.8	8365	21	AAK21445	Human defensin 2 p
32	61	67.8	209273	21	AAK21437	Human factor-relat
33	60	66.7	473	25	AB236740	Human GENSERT codin
34	59	65.6	411	14	ABO60009	Human colon cancer
35	59	65.6	424	24	AAQ53216	Sequence encoding
36	59	65.6	449	24	ABO60783	Human defensin 5 D
37	59	65.6	543	25	AB236767	Human GENSERT codin
38	59	65.6	1097	15	AAQ68911	Human c-myc far up
39	58	64.4	2880	17	AAQ53218	Genomic sequence e
40	58	64.4	92	17	AAK33499	92mer used in cons
41	58	64.4	92	18	AAK63062	Human pre-pro cart
42	58	64.4	92	19	AAV10898	DNA fragment SYN-U
43	56	62.2	105	19	AAK60903	Human cancer relat
44	55	61.1	537	24	ABN65503	Human LP321 DNA.
45	53	58.9	480	25	AAQ47375	

ALIGNMENTS

RESULT 1
AAK26697
AAK26697 standard; CDNA; 453 BP.

AAK26697:
18-JUN-1999 (first entry)

CDNA sequence encoding human defensin (Def-X) protein.

Human defensin, Def-X; antimicrobial; antiparasitic; pesticide;
cytotoxic; anticancer; inflammation; tissue repair;
endocrine regulation; corticostatic regulation; cancer; melanoma;
AIDS; immune deficiency; psoriasis; ss.

XX Homo sapiens.
OS

```
PN FR2767832-A1.
XX
XX 05-MAR-1999.
XX
XX 29-AUG-1997; 97FR-0010823.
XX
XX 29-AUG-1997; 97FR-0010823.
XX
XX 29-AUG-1997; 97FR-0010823.
XX
XX (GST ) GENSET SA.
XX
XX Bougueleret L, Chumakov I;
XX
XX WPI; 1999-183266/16.
XX
XX P-PSDB; AAX01604.
XX
XX Human defensin polypeptide Def-X - useful as antimicrobial agent,
XX anticancer agent, pesticide, etc.
XX
XX PS Disclosure; Fig 3; 56pp; French.
XX
XX The present sequence encodes human defensin (Def-X). The Def-X
XX polypeptide can be used as an antimicrobial, antiparasitic agent or
XX a pesticide. The Def-X polypeptide can be used as a cytostatic
XX (especially anticancer) agent, and as an agent for modulating processes
XX of inflammation, tissue repair and endocrine (especially corticostatic)
XX regulation. The polypeptide can be used in a composition for external
XX topical use, especially in a cosmetic composition. Compositions
XX containing Def-X can be used for prevention and treatment of microbial
XX and parasitic infections, especially where the microbial or parasitic
XX infections are Gram-positive or -negative bacterial infections or
XX mycobacterial, fungal or spirochaete infections, or where the viral
XX infections are associated with enveloped viruses, especially HSV and HIV.
XX The compositions can be used for prevention and/or treatment of cancers,
XX especially melanomas, or liver cancer, prostate cancer, non-small-cell
XX lung cancer or colorectal carcinoma, and for enhancing immunity,
XX especially in the case of AIDS, or preventing immune deficiency,
XX especially in the treatment of psoriasis, as well as for modulating
XX inflammatory processes, especially in the case of chronic inflammatory
XX disorders.
XX
XX SQ Sequence 453 BP; 107 A; 121 C; 97 G; 128 T; 0 other;

Alignment Scores:
Pred. No.: 1.09e-06 Length: 453
Score: 90.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-10-045-180A-4 (1-19) x AAX26697 (1-453)
QY
DB 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
52 ATGAGGAGCCCTCACCCTCTCTGCTTCTCTGCTGCTGAGGCTTCTGAGGCTGAGCA 108

RESULT 2
AAX26698
ID AAX26698 standard; cDNA; 453 BP.
XX
XX AAX26698;
XX
XX 18-JUN-1999 (first entry)
XX
XX cDNA sequence encoding human defensin (Def-X) protein.
XX
XX Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
XX cytostatic; anticancer; inflammation; tissue repair;
XX endocrine regulation; corticostatic regulation; cancer; melanoma;
XX AIDS; immune deficiency; psoriasis; ss.
XX
XX OS Homo sapiens.
XX
XX
```

```
PN FR2767832-A1.
XX
XX 05-MAR-1999.
XX
XX 29-AUG-1997; 97FR-0010823.
XX
XX 29-AUG-1997; 97FR-0010823.
XX
XX 29-AUG-1997; 97FR-0010823.
XX
XX (GST ) GENSET SA.
XX
XX Bougueleret L, Chumakov I;
XX
XX WPI; 1999-183266/16.
XX
XX Human defensin polypeptide Def-X - useful as antimicrobial agent,
XX anticancer agent, pesticide, etc.
XX
XX PS Claim 5; Fig 4; 56pp; French.
XX
XX The present sequence encodes human defensin (Def-X). The Def-X
XX polypeptide can be used as an antimicrobial, antiparasitic agent or
XX a pesticide. The Def-X polypeptide can be used as a cytostatic
XX (especially anticancer) agent, and as an agent for modulating processes
XX of inflammation, tissue repair and endocrine (especially corticostatic)
XX regulation. The polypeptide can be used in a composition for external
XX topical use, especially in a cosmetic composition. Compositions
XX containing Def-X can be used for prevention and treatment of microbial
XX and parasitic infections, especially where the microbial or parasitic
XX infections are Gram-positive or -negative bacterial infections or
XX mycobacterial, fungal or spirochaete infections, or where the viral
XX infections are associated with enveloped viruses, especially HSV and HIV.
XX The compositions can be used for prevention and/or treatment of cancers,
XX especially melanomas, or liver cancer, prostate cancer, non-small-cell
XX lung cancer or colorectal carcinoma, and for enhancing immunity,
XX especially in the case of AIDS, or preventing immune deficiency,
XX especially in the treatment of psoriasis, as well as for modulating
XX inflammatory processes, especially in the case of chronic inflammatory
XX disorders.
XX
XX SQ Sequence 453 BP; 107 A; 121 C; 97 G; 128 T; 0 other;

Alignment Scores:
Pred. No.: 1.09e-06 Length: 453
Score: 90.00 Matches: 19
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 20 Gaps: 0

US-10-045-180A-4 (1-19) x AAX26698 (1-453)
QY
DB 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
52 ATGAGGAGCCCTCACCCTCTCTGCTTCTCTGCTGCTGAGGCTTCTGAGGCTGAGCA 108

RESULT 3
AAX26696
ID AAX26696 standard; DNA; 4415 BP.
XX
XX AAX26696;
XX
XX 18-JUN-1999 (first entry)
XX
XX Nucleotide sequence of human defensin (Def-X).
XX
XX Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
XX cytostatic; anticancer; inflammation; tissue repair;
XX endocrine regulation; corticostatic regulation; cancer; melanoma;
XX AIDS; immune deficiency; psoriasis; ss.
XX
XX OS Homo sapiens.
XX
XX PN FR2767832-A1.
```


XX 05-MAR-1999.
XX 29-AUG-1997; 97FR-0010823.
XX 29-AUG-1997; 97FR-0010823.
XX (GEST) GENSET SA.
XX Bouquelerec L, Chumakov I;
XX WPI; 1999-183266/16.
XX Human defensin polypeptide Def-X - useful as antimicrobial agent,
XX anticancer agent, pesticide, etc.
XX Claim 5; Fig 2; 56pp; French.
XX The present sequence encodes human defensin (Def-X). The Def-X
XX polypeptide can be used as an antimicrobial, antiparasitic agent or
XX a pesticide. The Def-X polypeptide can be used as a cytostatic
XX (especially anticancer) agent, and as an agent for modulating processes
XX of inflammation, tissue repair and endocrine (especially corticostatic)
XX regulation. The polypeptide can be used in a composition for external
XX topical use, especially in a cosmetic composition. Compositions
XX containing Def-X can be used for prevention and treatment of microbial
XX and parasitic infections, especially where the microbial or parasitic
XX infections are Gram-positive or -negative bacterial infections or
XX mycobacterial, fungal or spirochaete infections, or where the viral
XX infections are associated with enveloped viruses, especially HSV and HIV.
XX The compositions can be used for prevention and/or treatment of cancers,
XX especially melanomas, or liver cancer, prostate cancer, non-small-cell
XX lung cancer or colorectal carcinoma, and for enhancing immunity,
XX especially in the case of AIDS, or preventing immune deficiency,
XX especially in the treatment of psoriasis, as well as for modulating
XX inflammatory processes, especially in the case of chronic inflammatory
XX disorders.
XX SQ Sequence 4415 BP; 1128 A; 1109 C; 912 G; 1252 T; 14 other;
XX
XX Alignment Scores:
XX Pred. No.: 1.79e-05 Length: 4415
XX Score: 90.00 Matches: 19
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: Gaps: 0
XX
XX US-10-045-180A-4 (1-19) x AAH57220 (1-4415)
XX QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
XX |||||
XX DB 3406 ATGAGAACCTCCACCTCTCTGCTTCTCTGAGGCGCTTCAAGCGCTGAGCA 3462
XX |||||
XX RESULT 4
XX AAH57220
XX ID AAH57220 standard; cDNA; 294 BP.
XX AC AAH57220;
XX XX
XX DT 10-SEP-2001 (first entry)
XX DE Human intestine specific cDNA sequence SEQ ID NO:60.
XX XX
XX KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
XX lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
XX metabolic disease; developmental disease; cytostatic; immunomodulatory;
XX neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX OS Homo sapiens.
XX XX
XX PN WO200132927-A2.
XX XX

PD 10-MAY-2001.
XX 02-NOV-2000; 2000MO-US30396.
XX 04-NOV-1999; 99US-0163508.
XX (INCY-) INCYTE GENOMICS INC.
XX Sornasse T, Sellhammer JJ, Watson GA;
XX WPI; 2001-291057/30.
XX New cell and tissue specific polynucleotides useful for diagnosis,
XX prognosis or monitoring of treatments for disorders where the gene is
XX associated with a cancer, immunopathology or neuropathology -
XX Claim 1; Page 95; 327pp; English.
XX AAH57561 to AAH57576 represent cell and tissue specific polynucleotide
XX sequences (I). (I) can have cytostatic, immunomodulatory and
XX neuroprotective activities, and can be used in gene therapy. (I) and
XX proteins (II) encoded by them are used in high throughput screening
XX assays to select DNA molecules, RNA molecules, peptide nucleic acids,
XX mimetics, peptides, proteins, agonists, antagonists, antibodies or
XX their fragments, immunoglobulins, inhibitors, drug compounds and
XX pharmaceutical agents. Expression of (I) in a sample indicates the
XX differentiation of embryonic stem cells into a tissue selected from
XX brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
XX tissues. (I) and (II) are used to produce an expression profile that
XX defines a metabolic or developmental process, treatment condition,
XX disease or disorder. The gene profile can be used for diagnosis,
XX prognosis or monitoring of treatments and for investigating a
XX predisposition to a disorder where the gene is associated with a
XX cancer, immunopathology or neuropathology.
XX SQ Sequence 294 BP; 72 A; 80 C; 73 G; 64 T; 5 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.0247 Length: 294
XX Score: 64.00 Matches: 15
XX Percent Similarity: 89.47% Conservative: 2
XX Best Local Similarity: 78.95% Mismatches: 2
XX Query Match: 71.11% Indels: 0
XX DB: Gaps: 0
XX
XX US-10-045-180A-4 (1-19) x AAH57220 (1-294)
XX QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
XX |||||
XX DB 15 ATGAGAACCTCCACCTCTCTGCTTCTCTGAGGCGCTTCAAGCGCT 71
XX |||||
XX RESULT 5
XX AHN59997
XX ID AHN59997 standard; cDNA; 398 BP.
XX AC AHN59997;
XX XX
XX DT 28-JUN-2002 (first entry)
XX DE Novel human coding sequence SEQ ID NO: 408.
XX XX
XX KW Human; antineoplastic; vulnary; antiinflammatory; immunomodulator;
XX antifertility; cerebroprotective; cytostatic; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag; gene; ss.
XX OS Homo sapiens.
XX XX
XX PN WO200222660-A2.
XX XX
XX PD 21-MAR-2002.
XX XX
XX PF 10-SEP-2001; 2001MO-US26015.
XX XX

```

XX PR 11-SEP-2000; 2000US-0659671.
XX
XX PA (HYSR-) HYSRQ INC.
XX
XX PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
XX PI Xue AJ, Yang Y, Weinman T, Drmanac RT;
XX
XX DR WPI; 2002-292408/33.
XX DR P-PSDB; ABB97584.
XX
XX PT An isolated polynucleotide for treating diseases associated with its
XX PT encoded polypeptide such as cancer and multiple sclerosis -
XX
XX PS Claim 1; SEQ ID NO 408; 509pp; English.
XX
XX CC The present invention provides the protein and coding sequences of 444
XX CC novel human proteins. These were isolated from expressed sequences tags
XX CC (ESTs). They can be used to stimulate cell growth, to regulate
XX CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
XX CC e.g. in burn treatment, to regulate the immune system e.g. to treat
XX CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
XX CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
XX CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
XX CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
XX CC Parkinson's disease. The present sequence is a coding sequence of the
XX CC invention.
XX
XX SQ Sequence 398 BP; 102 A; 108 C; 93 G; 95 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.0359 Length: 398
XX Score: 64.00 Matches: 15
XX Percent Similarity: 89.47% Conservative: 2
XX Best Local Similarity: 78.95% Mismatches: 2
XX Query Match: 71.11% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-10-045-180A-4 (1-19) x ABB95997 (1-398)
XX
XX QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
XX Db 30 ATGAGAACCTCCTACCTCCTCCTGCTGCTCTCTCTCTGAGCCCTCCAGGCCAAGGCT 86
XX
XX RESULT 6
XX ID AAQ53217 standard; cDNA; 452 BP.
XX
XX AC AAQ53217;
XX
XX DT 25-MAR-2003 (updated)
XX DT 17-JUN-1994 (first entry)
XX
XX DE Sequence encoding gastrointestinal defensin (GID) peptide called
XX DE human defensin 6.
XX
XX KW Gastrointestinal defensin peptide; GID; pharmaceutical; Paneth
XX KW cell; antimicrobial; anti-inflammatory; diagnosis;
XX KW contact disinfectant; ss.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT CDS 19..321
XX FT /*tag= a
XX
XX PN WO9324513-A1.
XX
XX PD 09-DEC-1993.
XX
XX PF 18-MAY-1993; 93WO-US04740.
XX
XX PR 22-MAY-1992; 92US-0888232.

```

```

XX PA (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
XX
XX PI Bevis CL, Jones DE;
XX
XX DR WPI; 1993-405719/50.
XX DR P-PSDB; AAR44819.
XX
XX PT Gastrointestinal defensin peptide(s) - useful as antimicrobial
XX PT and anti-inflammatory agents and for detecting gastrointestinal
XX PT disorders
XX
XX PS Claim 4; page 48-49; 97pp; English.
XX
XX CC A probe (D5' oligo) having the sequence given in AAQ53219 (claimed)
XX CC was used to screen a population of clones to identify novel defensin
XX CC defensin peptides. Pref. the libraries were human genomic and cDNA
XX CC libraries. Hybridisation and partial sequence analysis of the
XX CC identified clones contained previously characterised myeloid derived
XX CC defensin sequences as well as new defensin related sequences. Two
XX CC clones expressing new defensin related sequences were extensively
XX CC characterised and found to contain genes selectively expressed in
XX CC Paneth cells of the small intestine. These Paneth cell-derived
XX CC defensins are designated human defensin 5 and human defensin 6 and
XX CC are referred to as gastrointestinal defensin peptides. Comparison
XX CC of the deduced AA sequences of the defensin 5 cDNA with the
XX CC previously reported preprodefensin shows significant similarity.
XX CC The deduced AA sequence of defensin 6 cDNA has features similar to
XX CC defensin 5 and the previously reported preprodefensin.
XX CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX SQ Sequence 452 BP; 123 A; 118 C; 98 G; 113 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.0419 Length: 452
XX Score: 64.00 Matches: 15
XX Percent Similarity: 89.47% Conservative: 2
XX Best Local Similarity: 78.95% Mismatches: 2
XX Query Match: 71.11% Indels: 0
XX DB: 14 Gaps: 0
XX
XX US-10-045-180A-4 (1-19) x AAQ53217 (1-452)
XX
XX QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
XX Db 19 ATGAGAACCTCCTACCTCCTCCTGCTGCTCTCTCTGAGCCCTCCAGGCCAAGGCT 75
XX
XX RESULT 7
XX ID ACA56600 standard; cDNA; 452 BP.
XX
XX AC ACA56600;
XX
XX DT 06-JUN-2003 (first entry)
XX
XX DE Human signalling pathway polynucleotide probe SEQ ID NO 1198.
XX
XX KW Human; probe; ss; array element; Parkinson's disease;
XX KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
XX KW immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT CDS US6500938-B1.
XX
XX PD 31-DEC-2002.
XX
XX PF 30-JAN-1998; 98US-0016434.
XX
XX PR 30-JAN-1998; 98US-0016434.
XX
XX PA (INCY-) INCYTE GENOMICS INC.

```

Pt	Au-Young J, Seilhamer JJ;
Dx	WPI; 2003-352189/33.
Pt	Combination of polynucleotide probes, useful as array elements in a
Pt	microarray for monitoring the expression of a number of target
Pt	polynucleotides -
Pt	Claim 1, SEQ ID NO 1198; 65pp; English.
XX	
CC	The invention relates to a combination which, comprises a number of
CC	polynucleotide probes comprising a sequence selected from one of the 1490
CC	sequences mentioned in the specification. The combination is useful as an
CC	array element in a microarray for monitoring the expression of a number
CC	of target polynucleotides. The microarray is particularly useful in the
CC	diagnosis and treatment of cancer and immunopathology and neuropathology.
CC	The microarray is useful in diagnostics and treatment regimens, drug
CC	discovery and development, toxicological and carcinogenicity studies,
CC	forensics and pharmacogenomics. The microarray is also useful for
CC	monitoring progression of diseases and for developing sophisticated
CC	profiles for the effects of currently available therapeutic drugs. The
CC	combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC	and genomic fragments and in research and diagnostic applications. The
CC	array can detect changes in expression in a large number of genes coding
CC	for different signaling pathway populations which can be used to diagnose
CC	various diseases including cancer e.g. adenocarcinoma and leukemia,
CC	immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC	and Parkinson's disease. The present sequence represents a polynucleotide
CC	probe of the invention.
CC	Note: The sequence data for this patent did not form part of the printed
CC	specification but was obtained in electronic format directly from USPO
CC	at seqdata.uspo.gov/sequence.html?DocID=06500938B1.
CC	
XX	
SQ	Sequence 452 BP; 124 A; 118 C; 98 G; 112 T; 0 other;
	Alignment Scores:
	Pred. No.: 0.0419 Length: 452
	Score: 64.00 Matches: 15
	Percent Similarity: 89.47% Conservative: 2
	Best Local Similarity: 78..95% Mismatches: 0
	Query Match: 71.11% Indels: 2
	DB: 25 Gaps: 0
	US-10-045-180A-4 (1-19) x ACAS6600 (1-452)
Oy	1 MetarqThreutRhleUeUserAlaphelEuVaiAlaLengClnAlaTrrAla 19 ::: ::: ::: ::: :::
Db	19 ATGAGAACCTCACCAATCTCCTACGTCTTCCTGTTGGCCCTCAGGCCAAGGCT 75
RESULT 8	
AAT33496	
ID	AAT33496 standard; DNA; 110 BP.
XX	
AC	AAT33496;
XX	
DT	21-MAY-1997 (first entry)
DE	
XX	
XX	110mer used in construction of pre pro cartridge.
KM	Cationic peptide; CEMA; CEME; ceeropia; melittin; bacterial growth;
KW	antibiotic activity; permeable; bacterial outer membrane;
KW	lipopolysaccharide; fusion peptide; inhibition; endotoxaemia;
KW	sepsis associated disorder; septic shock; ss.
OS	
XX	Synthetic.
PN	WO9628559-A1.
PD	19-SEP-1996.
PF	13-MAR-1996; 96MO-IB00431.
PR	13-MAR-1995; 95US-0405234.

```

XX PA (VUBR-) UNIV BRITISH COLUMBIA.
XX PI
XX PI Brown MH, Hancock RW, Kelly N, Piers KL;
XX DR
XX WPI; 1996-433831/43.
XX
XX PT Recombinant production of antimicrobial, cationic peptide - produced
XX PT in bacterial host as fusion peptide with an anionic portion for
XX PT suppressing the anti-microbial activity of the cationic peptide
XX
XX PS Example 2; Page 19, 90pp; English.
XX
XX CC The sequences given in AAT733488-500 and AAT40195-204 are
XX CC oligonucleotides which were used in the construction of vectors used in
XX CC the method of the invention. The method comprises constructing an
XX CC expression vector containing DNA encoding the cationic peptide adjacent
XX CC to DNA encoding an anionic carrier peptide which inhibits bacterial
XX CC proteases. A bacterial host cell is transformed with the vector and
XX CC cultured to express a fusion peptide comprising the cationic and anionid
XX CC peptides. The biologically active cationic peptide is then recovered.
XX CC The vectors constructed using these sequences encode CBME, CBMA or HNP-
XX CC (human neutrophil peptide 1). CBMA contains the first 18 amino acids of
XX CC cecropin A and the last 8 residues of melittin. CBMA is derived from
XX CC CBME by changing the C-terminal amino acid of CBME to include two
XX CC additional lys residues. This modification improves the antibiotic
XX CC activity of the peptide against many bacterial species; two fold, as
XX CC well as substantially enhancing its ability to permeabilise bacterial
XX CC outer membranes and to bind to lipopolysaccharide. CBMA and CBME act
XX CC to permeabilise bacterial membranes breaking down the outer membrane
XX CC permeability barrier. They may be used in conjunction with antibiotics
XX CC to inhibit bacterial growth and to treat an endotoxaemia or sepsis
XX CC associated disorders e.g. septic shock.
XX
XX SQ Sequence 110 BP; 21 A; 38 C; 32 G; 19 T; 0 other;
XX
XX Alignment Scores:
XX Pred. No.: 0.025 Length: 110
XX Score: 61.00 Matches: 14
XX Percent Similarity: 84.21% Conservative: 2
XX Best Local Similarity: 73.68% Mismatches: 3
XX Query Match: 67.78% Indels: 0
XX Gaps: 0
XX
XX US-10-045-180A-4 (1-19) x AAT73496 (1-110)
XX
XX QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
XX ::::::::::::::::::::::::::::::::::::::::::::::::::::::
XX Db 5 ATGAGGACCTCTGCGCATCTTGTCTGCTCTCTCTGTTGGTGGCCCGAGGCCAGGCT 61
XX
XX RESULT 9
XX AAT63059
XX ID AAT63059 standard; DNA; 110 BP.
XX AC AAT63059;
XX
XX DT 25-MAR-2003 (updated)
XX DT 08-MAY-1997 (first entry)
XX
XX DE 110-mer used in construction of pre pro cartridge.
XX
XX KM HNP-1; CBME; CBMA; human neutrophil protein; fusion; insect defensin;
XX KM cecropin A; bee venom; melittin; antibiotic activity; S. typhimurium;
XX KM permeability; bacterial outer membrane; lipopolysaccharide;
XX KM antibiotic; E. coli; Pseudomonas aeruginosa; E. cloacae; S. aureus; ss.
XX
XX OS Synthetic.
XX
XX PN US5593866-A.
XX
XX PD 14-JAN-1997.
XX
XX PF 20-DEC-1995; 95US-0575052.

```


PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 4; SEQ ID NO 24062; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a single exon nucleic acid
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 187 BP; 37 A; 60 C; 54 G; 36 T; 0 other;

Alignment Scores:
Pred. No.: 0.0479 Length: 187
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 22 Gaps: 0

US-10-045-180A-4 (1-19) x ABA75757 (1-187)

Qy 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 13 ATGAGGACCTCGCCATCTTCTGCTCCATTCCTGCTGGTGGCCCTGACGCCAGGCT 69

RESULT 12

AAK50398
ID AAK50398 standard; DNA; 187 BP.
AC AAK50398;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 24955.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
FN WO200157276-A2.
XX
PD 09-AUG-2001.

30-JAN-2001; 2001WO-US00668.

04-FEB-2000; 2000US-0180312.

26-MAY-2000; 2000US-0207456.

30-JUN-2000; 2000US-0608408.

03-AUG-2000; 2000US-0632366.

21-SEP-2000; 2000US-0234687.

27-SEP-2000; 2000US-0236359.

04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488900/53.

Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -

PS Example 4; SEQ ID NO: 24955; 658pp + Sequence Listing; English.

CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.

SQ Sequence 187 BP; 37 A; 60 C; 54 G; 36 T; 0 other;

Alignment Scores:
Pred. No.: 0.0479 Length: 187
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 22 Gaps: 0

US-10-045-180A-4 (1-19) x AAK50398 (1-187)

Qy 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 13 ATGAGGACCTCGCCATCTTCTGCTCCATTCCTGCTGGTGGCCCTGACGCCAGGCT 69

RESULT 13

ABS23860
ID ABS23860 standard; DNA; 187 BP.
XX
AC ABS23860;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 23851.
XX
KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.

OS Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

XX 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-234687P.

XX 27-SEP-2000; 2000US-236359P.

XX 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2002-114183/15.

Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples -

PS Claim 4; SEQ ID No 23851; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of
CC probes; the novel set of probes which hybridise at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC and hyaline membrane disease. The present sequence is a single exon
CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIFO at
CC ftp.wifo.int/pub/published_pct_sequences.

XX SQ Sequence 187 BP; 37 A; 60 C; 54 G; 36 T; 0 other;

Alignment Scores:
Pred. No.: 0.0479 Length: 187
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 24 Gaps: 0

US-10-045-180A-4 (1-19) x ABS23860 (1-187)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 13 ATGAGGACCCCTGCCATCTTGTGTCATTCCTCTGTCGCCCTTGCAGGCCCGGCT 69

RESULT 14

AAT26106
ID AAT26106 standard; cDNA to mRNA; 299 BP.

XX AC AAT26106;

XX DT 18-OCT-1996 (first entry)

XX Human gene signature HUMG508342.

DE Gene signature; messenger RNA; mRNA; relative abundance; frequency;
XX human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.

OS Homo sapiens.

XX W09514772-A1.

XX 01-JUN-1995.

XX 11-NOV-1994; 94WO-JP01916.

XX 12-NOV-1993; 93JP-0355504.

XX (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.

XX Matsubara K, Okubo K;

XX WPI; 1995-206931/27.

XX Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

XX Claim 1; Page 2003; 2245pp; Japanese.

XX A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.

XX SQ Sequence 299 BP; 67 A; 87 C; 83 G; 61 T; 1 other;

Alignment Scores:
Pred. No.: 0.0853 Length: 299
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 16 Gaps: 0

US-10-045-180A-4 (1-19) x AAT26106 (1-299)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 18 ATGAGGACCCCTGCCATCTTGTGTCATTCCTCTGTCGCCCTTGCAGGCCCGGCT 74

RESULT 15

AAT20911
ID AAT20911 standard; DNA; 464 BP.

XX AC AAT20911;

XX DT 14-MAR-2001 (first entry)

XX Human defensin 3 polynucleotide fragment #2478.

DE Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
XX human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antidiabetic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.

XX Homo sapiens.

XX WO200062736-A2.

26-OCT-2000.
24-MAR-2000; 2000WO-US08020.
06-APR-1999; 99US-0127958.
(UYEC-) UNIV EAST CAROLINA.
(NYCE/) NYCE J W.
Nyce JW;
WPI; 2000-679539/66.
Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -
Disclosure; Page 218; 1592pp; English.
The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.
SQ Sequence 464 BP; 110 A; 128 C; 119 G; 107 T; 0 other;
Alignment Scores:
Pred. No.: 0.146 Length: 464
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 21 Gaps: 0
US-10-045-180A-4 (1-19) x AAF20911 (1-464)
Qy 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 90 ATGAGGACCTCGCCATCTTCCTGGTGGCCCTGAGGCCAGGCT 146

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 15:06:18 ; Search time 65.6915 Seconds
(without alignments)
963.776 Million cell updates/sec

Title: US-10-045-180A-4

Perfect score: 90

Sequence: 1 MRTLTLGSAFLVALQAWA 19

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2211978 seqs, 1666101734 residues

Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer_p2n.model -DEV=xl
-Q=/cgn2_1/USFTO spool_p/US10045180/runat 17122003 145026 12635/app query.fasta 1.860
-DB=Published Applications NA -QFMT=fastcap -SUFFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPTXT=0 -UNIT5=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR_MIN=0 -ALIGN=15 -MODS=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10045180 @cgn 1.174 @runat 17122003 145026 12635
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-Fgapop=6 -Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US05_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description

1	90	100.0	453	14	US-10-013-770-2	Sequence 2, Appli
2	90	100.0	453	14	US-10-045-180A-2	Sequence 2, Appli
3	90	100.0	4415	14	US-10-013-770-1	Sequence 1, Appli
4	90	100.0	4415	14	US-10-045-180A-1	Sequence 1, Appli
5	61	67.8	57	12	US-10-367-169-45	Sequence 45, Appli
6	61	67.8	57	13	US-10-076-802-45	Sequence 45, Appli
7	61	67.8	101	12	US-10-367-169-34	Sequence 34, Appli
8	61	67.8	101	13	US-10-076-802-34	Sequence 34, Appli
9	61	67.8	187	9	US-09-864-761-31864	Sequence 31864, A
10	61	67.8	555	13	US-10-252-157-223	Sequence 223, App
11	61	67.8	597	9	US-09-864-761-15342	Sequence 15342, A
12	61	67.8	726	10	US-09-816-828-16	Sequence 16, Appli
13	61	67.8	1348	15	US-10-125-237-78	Sequence 78, Appli
14	61	67.8	1348	15	US-10-105-891-78	Sequence 78, Appli
15	60	66.7	598	13	US-10-027-632-48280	Sequence 48280, A
16	60	66.7	598	13	US-10-027-632-48281	Sequence 48281, A
17	60	66.7	598	13	US-10-027-632-48282	Sequence 48282, A
18	60	66.7	598	13	US-10-027-632-48283	Sequence 48283, A
19	60	66.7	598	13	US-10-027-632-48284	Sequence 48284, A
20	60	66.7	598	14	US-10-027-632-48280	Sequence 48280, A
21	60	66.7	598	14	US-10-027-632-48281	Sequence 48281, A
22	60	66.7	598	14	US-10-027-632-48282	Sequence 48282, A
23	60	66.7	598	14	US-10-027-632-48283	Sequence 48283, A
24	60	66.7	598	14	US-10-027-632-48284	Sequence 48284, A
25	56	62.2	372	11	US-09-918-995-18938	Sequence 18938, A
26	54	60.0	436	15	US-10-171-581-252	Sequence 252, App
27	50	55.6	542	14	US-10-045-180A-8	Sequence 8, Appli
28	50	55.6	654	13	US-10-027-632-141529	Sequence 141529,
29	50	55.6	654	14	US-10-027-632-141529	Sequence 141529,
30	50	55.6	741	13	US-10-027-632-141528	Sequence 141528,
31	50	55.6	741	13	US-10-027-632-141530	Sequence 141530,
32	50	55.6	741	14	US-10-027-632-141528	Sequence 141528,
33	50	55.6	741	14	US-10-027-632-141530	Sequence 141530,
34	50	55.6	860	14	US-10-044-090-473	Sequence 473, App
35	50	55.6	4295	15	US-10-045-180A-7	Sequence 7, Appli
36	49	54.4	366	15	US-10-095-672A-4	Sequence 4, Appli
37	49	54.4	806	14	US-10-044-090-136	Sequence 136, App
38	49	54.4	1375	10	US-09-263-959-289	Sequence 289, App
39	49	54.4	684973	10	US-09-263-959-1	Sequence 1, Appli
40	48	53.3	1182	13	US-10-180-454-1	Sequence 1, Appli
41	48	53.3	1191	11	US-09-892-877-11	Sequence 11, Appli
42	48	53.3	1191	11	US-09-948-783-11	Sequence 11, Appli
43	48	53.3	127197	11	US-09-754-853A-1	Sequence 1, Appli
44	47	52.2	637	13	US-10-027-632-247501	Sequence 247501,
45	47	52.2	637	14	US-10-027-632-247501	Sequence 247501,

ALIGNMENTS

RESULT 1

US-10-013-770-2
; Sequence 2, Application US/10013770
; Publication No. US20020115151A1

; GENERAL INFORMATION:

; APPLICANT: GENSET SA

; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC

; COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; THERAPEUTIC APPLICATIONS

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson & Bear

; STREET: 550 West C Street

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/013,770


```

; FILING DATE: 10-Dec-2001
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/486,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Hart, Daniel
;
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
;
; INFORMATION FOR SEQ ID NO: 2:
;
; SEQUENCE CHARACTERISTICS:
;
; LENGTH: 453 BASE PAIRS
; TYPE: NUCLEOTIDE
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-013-770-2

```

```
DB: 14 Gaps: 0
US-10-045-180A-4 (1-19) x US-10-013-770-1 (1-4415)
QY 1 MetArThrieuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 3406 ATGAGGACCGCTCACCTCTCTGCTTTCTCTGCTGCGCCCTTCAGGCGCTGGCA 3462
RESULT 4
US-10-045-180A-1
; Sequence 1, Application US/10045180A
; Publication No. US20020182703A1
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Compositi
; FILE REFERENCE: GEN-100D1
; CURRENT APPLICATION NUMBER: US/10/045,180A
; CURRENT FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: US 09/486,580
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: PCT/FR98/01864
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: FR 97/10823
; PRIOR FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4415
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(4415)
; OTHER INFORMATION: Def-X genomic sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (85)..(85)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (143)..(143)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (670)..(670)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (970)..(970)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1111)..(1111)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1150)..(1150)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: CAAT signal
; LOCATION: (1711)..(1714)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: TATA signal
; LOCATION: (1758)..(1767)
; OTHER INFORMATION:
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1780)..(1780)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1836)..(1874)
; OTHER INFORMATION: Exon 1
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1875)..(1880)
; OTHER INFORMATION: splice donor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1974)..(1974)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2117)..(2117)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2133)..(2133)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2155)..(2335)
; OTHER INFORMATION: Alu insertion
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2186)..(2186)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2191)..(2191)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2367)..(2367)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2710)..(2780)
; OTHER INFORMATION: L1 fragment insertion
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3391)..(3393)
; OTHER INFORMATION: splice acceptor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3394)..(3577)
; OTHER INFORMATION: Exon 2
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3406)..(3408)
; OTHER INFORMATION: Translation initiation codon (ATG)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3578)..(3583)
; OTHER INFORMATION: splice donor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4123)..(4123)
; OTHER INFORMATION: n = a, c, g, or t.
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4161)..(4163)
; OTHER INFORMATION: splice acceptor site
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4164)..(4379)
; OTHER INFORMATION: Exon 3
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4274)..(4276)
; OTHER INFORMATION: Translation termination codon (TAA)
; FEATURE:
; NAME/KEY: polyA_signal
```



```
; SEQ ID NO 34
; LENGTH: 101
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Primer EE
US-10-076-802-34

Alignment Scores:
Pred. No.: 0.0058      Length: 101
Score: 61.00          Matches: 14
Percent Similarity: 84.21%      Conservative: 2
Best Local Similarity: 73.68%    Mismatches: 3
Query Match: 67.78%            Indels: 0
DB: 13                  Gaps: 0

US-10-045-180A-4 (1-19) x US-10-076-802-34 (1-101)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 77 ATGCGTACTCTGGCTATCTCTGAGCTATTCGCTGTGTCACCTGCGAGGCTCAAGCG 21

RESULT 9
US-09-864-761-31864
; Sequence 31864, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117

; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31864
; LENGTH: 187
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF233439.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6e+02
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 24
; OTHER INFORMATION: SWISSPROT HIT: P11479, EVALUE 7.00e-17
; OTHER INFORMATION: NT HIT: X52053.1, EVALUE 1.00e-101
; OTHER INFORMATION: EST_HUMAN HIT: AA321316.1, EVALUE 1.00e-101
US-09-864-761-31864

Alignment Scores:
Pred. No.: 0.0143      Length: 187
Score: 61.00          Matches: 14
Percent Similarity: 84.21%      Conservative: 2
Best Local Similarity: 73.68%    Mismatches: 3
Query Match: 67.78%            Indels: 0
DB: 9                  Gaps: 0

US-10-045-180A-4 (1-19) x US-09-864-761-31864 (1-187)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 13 ATGAGACCTCGCCATCTCTGCTGCCATTCCTGCTGGTGGCCCTGCAGGCCAGGCT 69

RESULT 10
US-10-252-157-223
; Sequence 223, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US
; CURRENT APPLICATION NUMBER: US/10/252,157
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 223
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 903338.12
US-10-252-157-223

Alignment Scores:
Pred. No.: 0.0568      Length: 555
Score: 61.00          Matches: 14
Percent Similarity: 84.21%      Conservative: 2
Best Local Similarity: 73.68%    Mismatches: 3
Query Match: 67.78%            Indels: 0
DB: 13                  Gaps: 0

US-10-045-180A-4 (1-19) x US-10-252-157-223 (1-555)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 145 ATGAGACCTCGCCATCTCTGCTGCCATTCCTGCTGGTGGCCCTGCAGGCCAGGCT 201

RESULT 11
US-09-864-761-15342
; Sequence 15342, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
```

```
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15342
; LENGTH: 597
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AF233439.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6e+02
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 18
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 24
US-09-864-761-15342

Alignment Scores:
Pred. No.: 0.0623 Length: 597
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 9 Gaps: 0

US-10-045-180A-4 (1-19) x US-09-864-761-15342 (1-597)
; TITLE OF INVENTION: No. US20020150898A1 Nucleic Acids and
; FILE REFERENCE: 791CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/125,237
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/552,929

RESULT 12
US-09-816-828-16
; Sequence 16, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Zhiwei
; APPLICANT: Zhao, Qing A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816,828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pc_FL_genes Version 2.0
; SEQ ID NO 16
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(522)
US-09-816-828-16

Alignment Scores:
Pred. No.: 0.0797 Length: 726
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 10 Gaps: 0

US-10-045-180A-4 (1-19) x US-09-816-828-16 (1-726)
; Sequence 78, Application US/10125237
; Publication No. US2003002329A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhang, Jie
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US2003002329A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/125,237
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/552,929

RESULT 13
US-10-125-237-78
; Sequence 78, Application US/10125237
; Publication No. US2003002329A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhang, Jie
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US2003002329A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/125,237
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/552,929

Oy 1 MetArgThrLeuThrLeuLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
Db 241 ATGAGGACCCCTGGCCATCTTCTGTCATTCCTGCTGGCTGCAGGCCAGGCT 297
```

```
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 78
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1143)
US-10-125-237-78

Alignment Scores:
Pred. No.: 0.174 Length: 1348
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 15 Gaps: 0

US-10-045-180A-4 (1-19) x US-10-125-237-78 (1-1348)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
Db 859 ATGAGGACCCCTGCCATCTTGTGTCATTCTCTGTTGGCCCTGCAGGCCCGAGGCT 915

RESULT 14
US-10-105-891-78
; Sequence 78, Application US/10105891
; Publication No. US20030073099A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhang, Jie
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030073099A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 791C1P2A
; CURRENT APPLICATION NUMBER: US/10/105,891
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 78
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1143)
US-10-105-891-78

Alignment Scores:
Pred. No.: 0.174 Length: 1348
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 15 Gaps: 0

US-10-045-180A-4 (1-19) x US-10-105-891-78 (1-1348)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
Db 859 ATGAGGACCCCTGCCATCTTGTGTCATTCTCTGTTGGCCCTGCAGGCCCGAGGCT 915

; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 78
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1143)
US-10-125-237-78

Alignment Scores:
Pred. No.: 0.174 Length: 1348
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 15 Gaps: 0

US-10-045-180A-4 (1-19) x US-10-125-237-78 (1-1348)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
Db 859 ATGAGGACCCCTGCCATCTTGTGTCATTCTCTGTTGGCCCTGCAGGCCCGAGGCT 915

RESULT 15
US-10-027-632-48280
; Sequence 48280, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 48280
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-48280

Alignment Scores:
Pred. No.: 0.0959 Length: 598
Score: 60.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 66.67% Indels: 0
DB: 13 Gaps: 0

US-10-045-180A-4 (1-19) x US-10-027-632-48280 (1-598)
QY 1 MetArgThrLeuThrLeuSerAlaPheLeuLeuValAlaLeuGlnAlaTrpAla 19
Db 402 ATGAGGACCCCTGCCATCTTGTGTCATTCTCTGTTGGCCCTGCAGGCCCGAGGCT 458

Search completed: December 17, 2003, 21:35:59
Job time : 67.6915 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 19 Seconds
(without alignments)
441.393 Million cell updates/sec

Title: US-10-045-180A-4
Perfect score: 90
Sequence: 1 MRTLTLASFLVLAQAWA 19

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10045180/runat 17122003 145024 12553/app query.fasta_1.860
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=xni -MINMATCH=0.1 -LOOPCI=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180@cgn 1 193 @runat 17122003 145024 12553 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	90	100.0	453	4	US-09-486-580A-2
2	90	100.0	4415	4	US-09-486-580A-1
3	64	71.1	452	1	US-08-158-189-6
4	64	71.1	452	4	US-09-016-434-1198
5	63	70.0	195	1	US-08-158-189-9
6	63	70.0	202	1	US-08-158-189-8
7	61	67.8	110	1	US-08-575-052-9
8	61	67.8	110	1	US-08-614-516A-9
9	61	67.8	110	1	US-08-770-557-9
10	61	67.8	110	1	US-08-405-234-9
11	61	67.8	203	1	US-08-158-189-10
12	61	67.8	203	1	US-08-158-189-11

13	61	67.8	218	1	US-08-158-189-38	Sequence 38, Appl
14	59	65.6	424	1	US-08-158-189-4	Sequence 4, Appl
C 15	59	65.6	1097	1	US-08-021-6080-5	Sequence 5, Appl
C 16	59	65.6	1097	1	PCT-US94-01782-5	Sequence 5, Appl
C 17	59	65.6	2880	1	US-08-158-189-1	Sequence 1, Appl
C 18	59	65.6	92	1	US-08-575-052-12	Sequence 12, Appl
C 19	58	64.4	92	1	US-08-614-516A-12	Sequence 12, Appl
C 20	58	64.4	92	1	US-08-770-557-12	Sequence 12, Appl
C 21	58	64.4	92	1	US-08-405-234-12	Sequence 12, Appl
C 22	56	62.2	105	3	US-09-042-071-44	Sequence 44, Appl
C 23	52	57.8	5687	2	US-08-380-403A-3	Sequence 3, Appl
C 24	52	57.8	5687	2	US-08-895-8100-3	Sequence 3, Appl
C 25	52	57.8	5687	4	US-08-486-013-66	Sequence 66, Appl
C 26	51	56.7	445	2	US-08-482-279-66	Sequence 66, Appl
27	51	56.7	445	2	US-08-342-268-66	Sequence 66, Appl
28	51	56.7	445	2	US-09-015-968-66	Sequence 66, Appl
29	51	56.7	445	3	US-09-397-386-66	Sequence 66, Appl
30	51	56.7	445	4	US-08-486-013-69	Sequence 69, Appl
31	51	56.7	2408	1	US-08-482-279-69	Sequence 69, Appl
32	51	56.7	2408	2	US-09-015-968-69	Sequence 69, Appl
33	51	56.7	2408	3	US-09-397-386-69	Sequence 69, Appl
34	51	56.7	2408	4	US-08-423-383-5	Sequence 5, Appl
35	51	56.7	2408	1	US-08-437-353A-5	Sequence 5, Appl
36	49	54.4	366	2	US-08-486-013-48	Sequence 48, Appl
37	49	54.4	377	2	US-08-482-268-48	Sequence 48, Appl
38	49	54.4	377	2	US-09-015-968-48	Sequence 48, Appl
39	49	54.4	377	3	US-09-397-386-48	Sequence 48, Appl
40	49	54.4	377	3	US-08-486-013-52	Sequence 52, Appl
41	49	54.4	388	1	US-08-482-279-52	Sequence 52, Appl
42	49	54.4	388	2		
43	49	54.4				
44	49	54.4				
45	49	54.4				

ALIGNMENTS

RESULT 1
US-09-486-580A-2
; Sequence 2, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 453 BASE PAIRS
; TYPE: NUCLEOTIDE
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: cDNA

```
; ORIGINAL SOURCE:  
; ORGANISM: Homo sapiens  
US-09-486-580A-2
```

Alignment Scores:

Pred. No.:	Length:
Score: 3.49e-07	453
Percent Similarity: 100.00%	Matches: 19
Best Local Similarity: 100.00%	Conservative: 0
Query Match: 100.00%	Mismatches: 0
DB: 100.00%	Indels: 0
Gaps: 4	Gaps: 0

US-10-045-180A-4 (1-19) x US-09-486-580A-2 (1-453)

Qy 1 MetArgThrLeuThrLeuSerAlaPheLeuValalaLeuGlnAlaTrpAla 19
Db 52 ATGGAGACCTCCCTCCTTCCTTCTCGGTGCCTTTTGAGGCCCTTCAGGCCTGGCA 108

RESULT 2

US-09-486-580A-1
Sequence 1, Application US/09486580A
Patent No. 6329340
GENERAL INFORMATION:
APPLICANT: GENSET SA
TITLE OF INVENTION: HUMAN DEFENSIN DER-X GENE AND DNAC
TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 550 West C Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/486,580A
FILING DATE: FEBRUARY 25, 2000
ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.064CI
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4415 BASE PAIRS
TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: DNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Exon 1
LOCATION: 1836..1874
FEATURE:
NAME/KEY: Exon 2
LOCATION: 3394..3577
FEATURE:
NAME/KEY: Exon 3
LOCATION: 4161..4380
FEATURE:
NAME/KEY: start CDS
LOCATION: 3406..3408
FEATURE:
NAME/KEY: stop CDS
LOCATION: 4276..4278
FEATURE:
NAME/KEY: polyadenylation site

```
; LOCATION: 4374..4379  
US-09-486-580A-1
```

Alignment Scores:

Pred. No.:	Length:
Score: 5.33e-06	4415
Percent Similarity: 100.00%	Matches: 19
Best Local Similarity: 100.00%	Conservative: 0
Query Match: 100.00%	Mismatches: 0
DB: 100.00%	Indels: 0
Gaps: 4	Gaps: 0

US-10-045-180A-4 (1-19) x US-09-486-580A-1 (1-4415)

Qy 1 MetArgThrLeuThrLeuSerAlaPheLeuValalaLeuGlnAlaTrpAla 19
Db 3406 ATGGAGACCCTCACCTCCCTCTCTGCTTCTCTGCTGCCTTCAGGCCTGGCA 3462

RESULT 3

US-08-158-189-6
Sequence 6, Application US/08158189
Patent No. 5641497
GENERAL INFORMATION:
APPLICANT: Bevins, Charles L.
TITLE OF INVENTION: Gastrointestinal Defensin Peptides,
TITLE OF INVENTION: cDNA Sequences, Methods for Production and Use Thereof
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5641497ris
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/158,189
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/888,232
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Johnson, Philip S.
REGISTRATION NUMBER: 27,200
REFERENCE/DOCKET NUMBER: CH-0219
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
FEATURE:
NAME/KEY: CDS
LOCATION: 19..321
US-08-158-189-6

Alignment Scores:

Pred. No.:	Length:
Score: 0.0117	452
Percent Similarity: 64.00	Matches: 15
Best Local Similarity: 89.47%	Conservative: 2
Query Match: 78.95%	Mismatches: 2
DB: 71.11%	Indels: 0
Gaps: 1	Gaps: 0

US-10-045-180A-4 (1-19) x US-08-158-189-6 (1-452)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
|||||
Db 19 ATGAGAACCTTCACCATCTCAGTGTCTCTCTGTGCGCCCTCCAGGCCAAGGCT 75

RESULT 4

US-09-016-434-1198
; Sequence 1198, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/016,434
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:

; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1198:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g181546
US-09-016-434-1198

Alignment Scores:
Pred. No.: 0.0117 Length: 452
Score: 64.00 Matches: 15
Percent Similarity: 89.47% Conservative: 2
Best Local Similarity: 78.95% Mismatches: 2
Query Match: 71.11% Indels: 0
DB: 4 Gaps: 0

US-10-045-180A-4 (1-19) x US-09-016-434-1198 (1-452)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
|||||
Db 19 ATGAGAACCTTCACCATCTCAGTGTCTCTCTGTGCGCCCTCCAGGCCAAGGCT 75

RESULT 5

US-08-158-189-9
; Sequence 9, Application US/08158189
; Patent No. 5641497
; GENERAL INFORMATION:
; APPLICANT: Bevins, Charles L.

; APPLICANT: Jones, Douglas E.
; TITLE OF INVENTION: Gastrointestinal Defensin Peptides.
; TITLE OF INVENTION: cDNA Sequences, Methods for Production and Use Thereof
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5641497ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,189
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,232
; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: Johnson, Philip S.
; REGISTRATION NUMBER: 27,200
; REFERENCE/DOCKET NUMBER: CH-0219
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 195 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
US-08-158-189-9

Alignment Scores:
Pred. No.: 0.00641 Length: 195
Score: 63.00 Matches: 15
Percent Similarity: 84.21% Conservative: 1
Best Local Similarity: 78.95% Mismatches: 3
Query Match: 70.00% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-4 (1-19) x US-08-158-189-9 (1-195)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
|||||
Db 70 ATGAGAACCTTCGCTCTGCTGTGTCATTCTCTGTGCGCCCTCCAGGCCAAGGCT 126

RESULT 6

US-08-158-189-8
; Sequence 8, Application US/08158189
; Patent No. 5641497
; GENERAL INFORMATION:
; APPLICANT: Bevins, Charles L.
; APPLICANT: Jones, Douglas E.
; TITLE OF INVENTION: Gastrointestinal Defensin Peptides,
; TITLE OF INVENTION: cDNA Sequences, Methods for Production and Use Thereof
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5641497ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA: US/08/158,189
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/888,232
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Johnson, Philip S.
REGISTRATION NUMBER: 27,200
REFERENCE/DOCKET NUMBER: CH-0219
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 202 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
US-08-158-189-8

Alignment Scores:
Pred. No.: 0.00669 Length: 202
Score: 63.00 Matches: 15
Percent Similarity: 84.21% Conservative: 1
Best Local Similarity: 78.95% Mismatches: 3
Query Match: 70.00% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-4 (1-19) x US-08-158-189-8 (1-202)

Qy 1 MetArgThrLeuThrLeuSerAlaPheLeuValalaLeuGlnAlaTrpAla 19
Db 77 ATGAGGACCTCGCTGCTTGGTGGCCATTCTCTGCTGGCCCTGCGAGGCCAGGCT 133

RESULT 7
US-08-575-052-9
Sequence 9, Application US/08575052
Patent No. 5593866
GENERAL INFORMATION:
APPLICANT: HANCOCK, ROBERT E.W.
APPLICANT: PIERS, KEVIN L.
APPLICANT: BROWN, MELISSA H.
TITLE OF INVENTION: CATIONIC PEPTIDES AND METHOD FOR PRODUCTION
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
CITY: LOS ANGELES
STATE: CALIFORNIA
COUNTRY: U.S.A.
ZIP: 90067
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/575,052
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/110,502
FILING DATE: 20-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: WETHERELL, JR., PH.D., JOHN R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: PD2823
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/455-5100
TELEFAX: 619/455-5110

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 2..110
US-08-575-052-9

Alignment Scores:
Pred. No.: 0.00721 Length: 110
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-4 (1-19) x US-08-575-052-9 (1-110)

Qy 1 MetArgThrLeuThrLeuSerAlaPheLeuValalaLeuGlnAlaTrpAla 19
Db 5 ATGAGGACCTCGCCATCTTGGTGGCCATTCTCTGCTGGCCCTGCGAGGCCAGGCT 61

RESULT 8
US-08-614-516A-9
Sequence 9, Application US/08614516A
Patent No. 5688767
GENERAL INFORMATION:
APPLICANT: HANCOCK, ROBERT E.W.
APPLICANT: PIERS, KEVIN L.
APPLICANT: BROWN, MELISSA H.
APPLICANT: KELLY, NIAMH
TITLE OF INVENTION: TREATMENT OF ENDOTOXIN-ASSOCIATED
TITLE OF INVENTION: DISORDERS WITH CATIONIC PEPTIDES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: FISH & RICHARDSON P.C.
STREET: 4225 Executive Square, Ste. 1400
CITY: La Jolla
STATE: CALIFORNIA
COUNTRY: U.S.A.
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/614,516A
FILING DATE: 13-MAR-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: HAILE, PH.D., LISA A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07420/010001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/578-5070
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 110 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 2..110
US-08-614-516A-9

Alignment Scores:

Pred. No.: 0.00721 Length: 110
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-4 (1-19) x US-08-614-516A-9 (1-110)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
||||| :|:|:|
Db 5 ATGAGACCCCTGCGCATCTTGTGCGCATTCCTCTGCTGGCCCTGCAGGCCAGGCT 61

RESULT 9

US-08-770-557-9
; Sequence 9, Application US/08770557
; Patent No. 5707855
; GENERAL INFORMATION:
; APPLICANT: HANCOCK, ROBERT E.W.
; APPLICANT: PIERS, KEVIN L.
; APPLICANT: BROWN, MELISSA H.
; TITLE OF INVENTION: CATIONIC PEPTIDES AND METHOD FOR PRODUCTION
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SPENSLEY HORN JURAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 90067

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,557
; FILING DATE: 20-DEC-1996

CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/575,052
; FILING DATE:
; APPLICATION NUMBER: US/08/110,502
; FILING DATE: 20-AUG-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WETHERELL, JR., PH.D., JOHN R.
; REGISTRATION NUMBER: 31,678
; REFERENCE/DOCKET NUMBER: PD2823

TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:

NAME/KEY: CDS
; LOCATION: 2..110
US-08-770-557-9

Alignment Scores:

Pred. No.: 0.00721 Length: 110
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-4 (1-19) x US-08-770-557-9 (1-110)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
||||| :|:|:|
Db 5 ATGAGACCCCTGCGCATCTTGTGCGCATTCCTCTGCTGGCCCTGCAGGCCAGGCT 61

RESULT 10

US-08-405-234-9
; Sequence 9, Application US/08405234
; Patent No. 5789377
; GENERAL INFORMATION:
; APPLICANT: HANCOCK, ROBERT E.W.
; APPLICANT: PIERS, KEVIN L.
; APPLICANT: BROWN, MELISSA H.
; APPLICANT: KELLY, NIMAH
; TITLE OF INVENTION: TREATMENT OF ENDOTOXIN-ASSOCIATED DISORDERS
; TITLE OF INVENTION: WITH CATIONIC PEPTIDES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SPENSLEY HORN JURAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: U.S.A.
; ZIP: 90067

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,234
; FILING DATE: 13-MAR-1995

CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: PD3535
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110

INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 110 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..110
US-08-405-234-9

Alignment Scores:

Pred. No.: 0.00721 Length: 110
Score: 61.00 Matches: 14
Percent Similarity: 84.21% Conservative: 2
Best Local Similarity: 73.68% Mismatches: 3
Query Match: 67.78% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-4 (1-19) x US-08-405-234-9 (1-110)

QY 1 MetArgThrLeuThrLeuSerAlaPheLeuValAlaLeuGlnAlaTrpAla 19
||||| :|:|:|
Db 5 ATGAGACCCCTGCGCATCTTGTGCGCATTCCTCTGCTGGCCCTGCAGGCCAGGCT 61

RESULT 11

US-08-158-189-10
; Sequence 10, Application US/08158189
; Patent No. 5641497
; GENERAL INFORMATION:
; APPLICANT: Bevins, Charles L.
; APPLICANT: Jones, Douglas E.

Search completed: December 17, 2003, 14:54:00
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 1258.91 Seconds
(without alignments)
1429.821 Million cell updates/sec

Title: US-10-045-180A-5
Perfect score: 233
Sequence: 1 EPLQARAHENPAQKPPADD.....IYFGDDSCSLQVPGSTKGL 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2 i/USPTO.spool.p/US10045180/runat 17122003 145023 12528/app query.fasta_1.860
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180 @CGN 1.1 5283 @runat 17122003 145023 12528 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.em:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
15: em.ba:*
16: em.fun:*
17: em.hum:*
18: em.in:*
19: em.mu:*
20: em.om:*
21: em.or:*
22: em.ov:*
23: em.pat:*
24: em.ph:*
25: em.pl:*
26: em.ro:*
27: em.sts:*
28: em.un:*

RESULT 1

ALIGNMENTS

Result No.	Score	Query Match %	Length	DB ID	Description
1	233	100.0	453	6	A98571 Sequence 2
2	233	100.0	453	6	BD074746 Human def
3	209	89.7	4415	6	A98570 Sequence 1
4	209	89.7	4415	6	BD074745 Human def
5	201	86.3	128544	9	AF238378 Homo sapi
6	181	77.7	163093	2	AC116559 Papio anu
7	181	77.7	188604	2	AC116558 Papio anu
8	114	48.9	542	9	HSCCSP4
9	108	46.4	4580	9	AF195487
10	108	46.4	79730	9	AF134060 Homo sapi
11	108	46.4	150940	9	AF233439 Homo sapi
12	108	46.4	158653	9	AF200455 Homo sapi
13	108	46.4	168702	9	AC079018 Homo sapi
14	108	46.4	169044	9	AF205406 Homo sapi
15	104	44.6	496	9	AF355799 Homo sapi
16	101	43.3	452	6	AR270635 Sequence 6
17	101	43.3	452	6	I49588 Sequence 6
18	101	43.3	452	9	HUMDFNSIX
19	101	43.3	38845	9	AF314059 Homo sapi
20	98	42.1	3208	9	HSU10287
21	95	40.8	3377	9	AF104257
22	95	40.8	5124	9	HUNPIV
23	90	38.6	398	6	AX405993 Sequence
24	90	38.6	3060	9	HSU33317
25	90	38.6	232180	2	AC021883
26	86	36.9	243	6	AR181803
27	86	36.9	243	6	AX047114
28	86	36.9	421	9	AF184157
29	84.5	36.3	469	4	RABDNP3AB
30	83	35.6	460	9	AF184156
31	83	35.6	495	6	AR181796
32	83	35.6	495	6	AX047101
33	83	35.6	495	9	AF191101
34	83	35.6	3090	4	RABDNP3AA
35	82	35.2	421	9	AF184158
36	82	35.2	500	6	AR181795
37	82	35.2	500	6	AX047099
38	82	35.2	500	9	AF191100
39	80	34.3	424	6	I49587
40	80	34.3	2548	6	AX047111
41	80	34.3	2548	9	AF191103
42	80	34.3	2553	6	AR181800
43	78.5	33.7	448	9	HUMDEF1A
44	78.5	33.7	451	9	HUMDEFB
45	78.5	33.7	464	9	HSDEFF3

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

A98571
LOCUS A98571 453 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 2 from Patent WO9911663.
ACCESSION A98571
VERSION A98571.1 GI:6781627
KEYWORDS unidentified
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 453)
AUTHORS Bougueleret,L. and Chumakov,I.
TITLE HUMAN DEFENSIN DEF-X, GENE AND DNAC, COMPOSITION CONTAINING SAME
AND DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
JOURNAL Patent: WO 9911663-A 2 11-MAR-1999;
BOUGUELERET LYDIE (FR); CHUMAKOV ILYA (FR)
FEATURES
source Location/Qualifiers
1..453
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 107 a 121 c 97 g 128 t
ORIGIN
Alignment Scores:
Pred. No.: 1.7e-24 Length: 453
Score: 233.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-045-180A-5 (1-44) x A98571 (1-453)
Qy 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAsp 20
Db 109 GAGCGCTCCAGGCAAGAGCTCATGAGATGCCAGCCAGAGCGCTCCAGCAGATGAC 168
Qy 21 GlnAspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySer 40
Db 169 CAGGATGTGTCATTACTTTTCAGGAGATGACAGCTGCTCTCTTCAGGTTCCAGGCTCA 228
Qy 41 ThrLysGlyLeu 44
Db 229 ACAAGGGCTTG 240
RESULT 2
BD074746
LOCUS BD074746 453 bp DNA linear PAT 27-AUG-2002
DEFINITION Human defensin polypeptide Def-X, genome DNA and cDNA, composition containing the same, and application to diagnosis and remedy treatment.
ACCESSION BD074746
VERSION BD074746.1 GI:22620349
KEYWORDS JP 2001514264-A/2.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 453)
AUTHORS Bougueleret,L. and Shmacov,I.
TITLE Human defensin polypeptide Def-X, genome DNA and cDNA, composition containing the same, and application to diagnosis and remedy
JOURNAL Patent: JP 2001514264-A 2 11-SEP-2001;
COMMENT GENSET
OS Homo sapiens (human)
PN JP 2001514264-A/2
PD 11-SEP-2001
PF 28-AUG-1998 JP 2000508701
PR 29-AUG-1997 FR 97/10823
PI LYDIE BOUGUELERET,ILYA SHMACOV
PC C07K14/435,A01N43/50,A01N63/00,A61K7/00,A61K38/00,A61P23/00,
PC A61P35/00,
PC A61P37/02,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N15/09, PC

C12P21/02,
PC C12P21/08,C12Q1/68,G01N33/53,A61K37/02,C12N15/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Human defensin polypeptide Def-X, genome DNA and cDNA, CC
composition
CC containing the same, and application to diagnosis and remedy
CC treatment
FH Key Location/Qualifiers
FT source 1..453
FT /organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1..453
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 107 a 121 c 97 g 128 t
ORIGIN
Alignment Scores:
Pred. No.: 1.7e-24 Length: 453
Score: 233.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0
US-10-045-180A-5 (1-44) x BD074746 (1-453)
Qy 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAsp 20
Db 109 GAGCGCTCCAGGCAAGAGCTCATGAGATGCCAGCCAGAGCGCTCCAGCAGATGAC 168
Qy 21 GlnAspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySer 40
Db 169 CAGGATGTGTCATTACTTTTCAGGAGATGACAGCTGCTCTCTTCAGGTTCCAGGCTCA 228
Qy 41 ThrLysGlyLeu 44
Db 229 ACAAGGGCTTG 240
RESULT 3
A98570
LOCUS A98570 4415 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 1 from Patent WO9911663.
ACCESSION A98570
VERSION A98570.1 GI:6781626
KEYWORDS unidentified
SOURCE unclassified.
REFERENCE 1 (bases 1 to 4415)
AUTHORS Bougueleret,L. and Chumakov,I.
TITLE HUMAN DEFENSIN DEF-X, GENE AND DNAC, COMPOSITION CONTAINING SAME
AND DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
JOURNAL Patent: WO 9911663-A 1 11-MAR-1999;
BOUGUELERET LYDIE (FR); CHUMAKOV ILYA (FR)
FEATURES
source Location/Qualifiers
1..4415
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
exon 1836..1874
exon 3394..3577
exon 4161..4380
PolyA_site 4374..4379
BASE COUNT 1128 a 1109 c 912 g 1252 t 14 others
ORIGIN
Alignment Scores:
Pred. No.: 6.65e-20 Length: 4415
Score: 209.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0


```

Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 89.70%  Indels: 0
DB: 6  Gaps: 0

US-10-045-180A-5 (1-44) x A98570 (1-4415)

QY 1 GluProLeuGlnAlaAargAlaHisGluMetProAlaGlnLysGlnProAlaAsp 20
Db 3463 GAGCCGCTCCAGCAAGAGCTCATGAGATGCCAGCCAGAGCAGCTCCAGCATGAC 3522

US-10-045-180A-5 (1-44) x A98570 (1-4415)

QY 21 GlnAspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGly 39
Db 3523 CAGGATGGTTCATTACTTTTCAGGAGATGACAGCTGCTCTCTTCAGTTCCAGGT 3579

RESULT 4
LOCUS BD074745 4415 bp DNA linear PAT 27-AUG-2002
DEFINITION Human defensin polypeptide Def-X, genome DNA and cDNA, composition
containing the same, and application to diagnosis and remedy
treatment.
ACCESSION BD074745
VERSION BD074745.1 GI:22620348
KEYWORDS JP 2001514264-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 4415)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Bougueleret, L. and Shmacov, I.
AUTHORS Human defensin polypeptide Def-X, genome DNA and cDNA, composition
containing the same, and application to diagnosis and remedy
PATENT: JP 2001514264-A 1 11-SEP-2001;
JOURNAL GENSET
COMMENT OS Homo sapiens (human)
PN JP 2001514264-A/1
PD 11-SEP-2001
PF 28-AUG-1998 JP 2000508701
PI 29-AUG-1997 FR 97/10823
PR LYDIE BOUGUELERET, ILYA SHMACOV
PC C07K14/435, A01N43/50, A01N63/00, A61K7/00, A61K38/00, A61P29/00,
A61P35/00,
PC A61P37/02, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N15/09, PC
C12P21/02,
PC C12P21/08, C12Q1/68, G01N33/53, A61K37/02, C12N15/00 CC
Strandedness: Double;
CC Topology: Linear;
CC Human defensin polypeptide Def-X, genome DNA and cDNA, CC
composition
CC containing the same, and application to diagnosis and remedy
treatment
FH Key Location/Qualifiers
FT exon 1836..1874
FT exon 3394..3577
FT exon 4161..4380
FT CDS 3406..3408
FT CDS 4276..4278.

FEATURES
source
1..4415
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

BASE COUNT 1128 a 1109 c 912 g 1252 t 14 others
ORIGIN

Alignment Scores:
Pred. No.: 6,65e-20 Length: 4415
Score: 209.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.70% Indels: 0
DB: 6 Gaps: 0

US-10-045-180A-5 (1-44) x BD074745 (1-4415)

QY 1 GluProLeuGlnAlaAargAlaHisGluMetProAlaGlnLysGlnProAlaAsp 20
Db 3463 GAGCCGCTCCAGCAAGAGCTCATGAGATGCCAGCCAGAGCAGCTCCAGCATGAC 3522

US-10-045-180A-5 (1-44) x A98570 (1-4415)

QY 21 GlnAspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGly 39
Db 3523 CAGGATGGTTCATTACTTTTCAGGAGATGACAGCTGCTCTCTTCAGTTCCAGGT 3579

RESULT 5
LOCUS AF238378/c 128544 bp DNA linear PRI 02-APR-2003
DEFINITION Homo sapiens chromosome 8 clone SCB-561b17 map p22-p21, complete
sequence.
ACCESSION AF238378
VERSION AF238378.5 GI:29469504
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 128544)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Rosenthal, A.
AUTHORS Chromosome 8 genomic sequence
Unpublished
REFERENCE 2 (bases 1 to 128544)
Schudy, A., Schilhabel, M., Schutte, B., Ganz, T., Linzmeier, R.,
Ho, C.H., Hoang, B.V., McCray, P., Baumgart, C., Menzel, U.,
Schattevoy, R. and Rosenthal, A.
Direct Submission
Submitted (22-FEB-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany
3 (bases 1 to 128544)
Genome Sequencing Center Jena.
Direct Submission
Submitted (24-MAY-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
4 (bases 1 to 128544)
Genome Sequencing Center Jena.
Direct Submission
Submitted (10-NOV-2000) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
5 (bases 1 to 128544)
Lagemann, D. and Platzter, M.
Direct Submission
Submitted (06-JUL-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
6 (bases 1 to 128544)
Lagemann, D. and Platzter, M.
Direct Submission
Submitted (02-APR-2003) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
On Apr 2, 2003 this sequence version replaced gi:21700555.
----- Genome Center
Center: Insitute of Molecular Biotechnology
Center code: IMB
Web site: http://genome.imb-jena.de/
Contact: gacj-submit@genome.imb-jena.de
----- Project Information
Center project name: H370
Center clone name: SCB-561b17
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 127999 bases at least Q40
Consensus quality: 128330 bases at least Q30
Consensus quality: 128544 bases at least Q20
Quality coverage: 11.05x
-----

```

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

----- Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

Location/Qualifiers

```

FEATURES
    source
        1..128544
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="8"
            /map="p22-p21"
            /clone="SCb-561b17"
            44529..44595
                /note="single stranded/single chemistry region"
            45703
                /note="low quality region"
            45815..45872
                /note="low quality region"
            45824..45872
                /note="single stranded/single chemistry region"
            45843
                /note="low quality region"
            45847
                /note="low quality region"
            45849
                /note="low quality region"
            45858
                /note="low quality region"
            45868
                /note="low quality region"
            46235
                /note="low quality region"
            47899..47900
                /note="low quality region"
            47899..47900
                /note="low quality region"
            49172
                /note="low quality region"
            49172
                /note="low quality region"
            49220
                /note="low quality region"
            49339
                /note="low quality region"
            49756..49807
                /note="single stranded/single chemistry region"
            49961
                /note="low quality region"
            52933..52948
                /note="low quality region"
            52943
                /note="low quality region"
            53544
                /note="low quality region"
            53612..53926
                /note="single stranded/single chemistry region"
            53674..53677
                /note="low quality region"
            53779..53799
                /note="low quality region"
            53787
                /note="low quality region"
            54734..55129
                /note="single stranded/single chemistry region"
            56028..56165
                /note="single stranded/single chemistry region"
            57952..58060
                /note="single stranded/single chemistry region"

```

```

misc_feature 57968..57974
             /note="low quality region"
unsure 57968
unsure 57970
misc_feature 57974
             /note="low quality region"
misc_feature 58652
             /note="low quality region"
misc_feature 59561..59638
             /note="single stranded/single chemistry region"
misc_feature 59925
             /note="low quality region"
misc_feature 59925
             /note="low quality region"
misc_feature 59929
             /note="low quality region"
misc_feature 61695..62103
             /note="single stranded/single chemistry region"
misc_feature 61858..61859
             /note="low quality region"
misc_feature 62263..62485
             /note="single stranded/single chemistry region"
misc_feature 62655..63118
             /note="single stranded/single chemistry region"
misc_feature 62750..62785
             /note="low quality region"
unsure 62774..62777
unsure 62783
misc_feature 62801..62825
             /note="low quality region"
unsure 62819
unsure 62822
misc_feature 62872..62874
             /note="low quality region"
misc_feature 62908
             /note="low quality region"
misc_feature 62917..62987
             /note="low quality region"
unsure 62933..62935
unsure 62955
unsure 62957
misc_feature 63082
             /note="low quality region"
misc_feature 63110
             /note="low quality region"
misc_feature 63119..63281
             /note="low quality region"
misc_feature 63119..63141
             /note="single clone coverage"
misc_feature 63129..63131
             /note="single stranded/single chemistry region"
misc_feature 63142..63258
             /note="single stranded/single chemistry region"
unsure 63159
unsure 63197..63258
unsure 63281
misc_feature 63876..63878
             /note="low quality region"
misc_feature 63933..64021
             /note="single stranded/single chemistry region"
misc_feature 63943
             /note="low quality region"
misc_feature 63984..63985
             /note="low quality region"
misc_feature 64004..64007
             /note="low quality region"
misc_feature 70360
             /note="low quality region"

```

Alignment Scores:

```

Pred. No.: 4,33e-17 Length: 128544
Score: 201.00 Matches: 37
Percent Similarity: 97.44% Conservative: 1
Best Local Similarity: 94.87% Mismatches: 1
Query Match: 86.27% Indels: 0
DB: 9 Gaps: 0

```

```

US-10-045-180A-5 (1-44) x AF238378 (1-128544)
QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAsp 20
Db 63266 GAGCGCTCCGCGAAGAGCTCATGAGATGCCAGCCAGAGCAGCTCCAGCAGATGAC 63207

QY 21 GlnAspValValIleTyPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 63206 CAGGATGTGGTCTTACTTTTTCAGGAGATGACAGCTGCTCTCTTCAGGTTCCAGGT 63150

RESULT 6
AC116559 163093 bp DNA linear HTG 30-MAY-2003
LOCUS Papio anubis clone rp41-339c10, WORKING DRAFT SEQUENCE, 3 ordered
DEFINITION pieces.
ACCESSION AC116559
VERSION AC116559.18 GI:31193968
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Papio.
1 (bases 1 to 163093)
XU, W., Hua, A., Eichler, E. and Roe, B.A.
Papio anubis BAC Clone rp41-339c10
Unpublished
2 (bases 1 to 163093)
XU, W., Hua, A., Eichler, E. and Roe, B.A.
Direct Submission
Submitted (30-MAY-2003) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 163093)
XU, W., Hua, A., Eichler, E. and Roe, B.A.
Direct Submission
Submitted (30-MAY-2003) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 30, 2003 this sequence version replaced gi:30725962.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma
Center code:UOKNOR
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 6315: contig of 6315 bp in length
* 6316 6415: gap of unknown length
* 6416 16704: contig of 10289 bp in length
* 16705 16804: gap of unknown length
* 16805 163093: contig of 146289 bp in length.
* Location/Qualifiers
* 1. .163093
* /organism="Papio anubis"
* /mol_type="genomic DNA"
* /db_xref="taxon:9555"
* /clone="rp41-339c10"
* /clone_lib="RP41 - 41 Male (Olive) Baboon BAC Library"
BASE COUNT 44435 a 36508 c 35861 g 46065 t 224 others
ORIGIN

Alignment Scores:
Pred. No.: 4.44e-14 Length: 163093
Score: 181.00 Matches: 35

US-10-045-180A-5 (1-44) x AC116559 (1-163093)
QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAsp 20
Db 32923 GAGCGCTCCGCGAAGAGCTCATGAGATGCCAGCCAGAGCAGCTCCAGCAACGAC 32982

QY 21 GlnAspValValIleTyPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 32983 CAAGATGTGGTCTTACTTTTTCAGGAGATGACAGCTCTCTCTTCAGGTTCCAGGT 33039

RESULT 7
AC116558/c
LOCUS Papio anubis clone rp41-273g19, WORKING DRAFT SEQUENCE, 2 ordered
DEFINITION pieces.
ACCESSION AC116558
VERSION AC116558.16 GI:30725961
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE Papio anubis (olive baboon)
ORGANISM Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Papio.
1 (bases 1 to 188604)
XU, W., Hua, A., Eichler, E. and Roe, B.A.
Papio anubis BAC Clone rp41-273g19
Unpublished
2 (bases 1 to 188604)
XU, W., Hua, A., Eichler, E. and Roe, B.A.
Direct Submission
Submitted (29-MAR-2002) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 188604)
XU, W., Hua, A., Eichler, E. and Roe, B.A.
Direct Submission
Submitted (11-JUN-2003) Department of Chemistry And Biochemistry,
The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 15, 2003 this sequence version replaced gi:28173120.
----- Genome Center
Center: Department of Chemistry And Biochemistry
The University of Oklahoma
Center code:UOKNOR
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 6057: contig of 6057 bp in length
* 6058 6157: gap of unknown length
* 6158 188604: contig of 182447 bp in length.
* Location/Qualifiers
* 1. .188604
* /organism="Papio anubis"
* /mol_type="genomic DNA"
* /db_xref="taxon:9555"
* /clone="rp41-273g19"
* /clone_lib="RP41 - 41 Male (Olive) Baboon BAC Library"
BASE COUNT 53384 a 41769 c 40796 g 52555 t 100 others
ORIGIN

Alignment Scores:

```

```

Pred. No.: 5 23e-14 Length: 188604
Score: 181.00 Matches: 35
Percent Similarity: 94.87% Conservative: 2
Best Local Similarity: 89.74% Mismatches: 2
Query Match: 77.68% Indels: 0
DB: 2 Gaps: 0

US-10-045-180A-5 (1-44) x AC116558 (1-188604)

Qy 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAsp 20
Db 142609 GAGCCGCTCCAGGCAAGAGCTGATGAGTCCAGCCAGGAGAGCTCCAGCAACAGCAG 142550

Qy 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 142549 CAAGATGTGTGCTACTTTTCAGGAGATGACAGCTCTCTCTTCAGGTTCCAGGT 142493

RESULT 8
HSCCSHP4
LOCUS HSCCSHP4 542 bp mRNA linear PRI 07-JUN-1999
DEFINITION H.sapiens mRNA for corticostatin HP-4 precursor.
ACCESSION X65977
VERSION X65977.1 GI:29734
KEYWORDS corticostatin/defensin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 542)
AUTHORS Palfree,R.G., Sadro,L.C. and Solomon,S.
TITLE The gene encoding the human corticostatin HP-4 precursor contains a
recent 86-base duplication and is located on chromosome 8
JOURNAL Mol. Endocrinol. 7 (2), 199-205 (1993)
MEDLINE 93225961
PUBMED 8469233
REFERENCE 2 (bases 1 to 542)
AUTHORS Palfree,R.G.E.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1992) R.G.E. Palfree, McGill University Royal
Victoria Hosp., Rm L2.05, 687 Pine Ave West, Montreal PQ, H3A 1A1,
CANADA

FEATURES
source
1..542
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="8"
/clones="HP-4.1"
/tissue_type="bone marrow"
/clone_lib="Clontech HL1058a, human bone marrow cDNA"
/dev_stages="adult"
gene
1..542
/gene="HP-4"
CDS
52..345
/gene="HP-4"
/codon_start=1
/product="corticostatin/defensin HP-4 precursor"
/protein_id="CAA46792.1"
/db_xref="GI:29735"
/db_xref="SWISS-PROT:P12838"
/translation="MRIILALAILLVALQVRAGFLQARGDEAPQEQRGPEQDQIS
SFANDKSSALQVSSGRGMVCSRLVFCRRTELRVGNLIGVSVFTYCCTFVD"
sig_peptide
52..108
/gene="HP-4"
mat_peptide
109..342
/gene="HP-4"
/misc_feature
330..412
/product="corticostatin/defensin HP-4"
/note="recent duplication of immediate upstream segment"
BASE COUNT 118 a 137 c 135 g 152 t
ORIGIN

Alignment Scores:
Pred. No.: 3 46e-07 Length: 542
Score: 114.00 Matches: 23
Percent Similarity: 69.77% Conservative: 7
Best Local Similarity: 53.49% Mismatches: 13
Query Match: 48.93% Indels: 0
DB: 9 Gaps: 0

US-10-045-180A-5 (1-44) x HSCCSHP4 (1-542)

Qy 2 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspGln 21
Db 112 CCACCTCCAGGCAAGAGGTGATGAGGCTCCAGGCCAGGAGCGTGGCCAGAGACGAC 171

Qy 22 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 41
Db 172 GACATATCTATTTCCTTTTCATGGGATAAAAGCTGTGCTCTTCAGGTTTCAGGCTCAACA 231

Qy 42 LysGlyLeu 44
Db 232 AGGGGCGATG 240

RESULT 9
AF195487
LOCUS AF195487 4580 bp DNA linear PRI 09-DEC-1999
DEFINITION Homo sapiens defensin-related pseudogene.
ACCESSION AF195487
VERSION AF195487.1 GI:6540006
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4580)
AUTHORS Bevins,C.L. and Jones,D.B.
TITLE Analysis of the Human Defensin Gene Locus: Sequence of Clone HG27B
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4580)
AUTHORS Bevins,C.L. and Jones,D.B.
TITLE Direct Submission
JOURNAL Submitted (15-OCT-1999) Immunology, The Cleveland Clinic
Foundation, 9500 Euclid Avenue, Cleveland, OH 44195, USA

FEATURES
Location/Qualifiers
source
1..4580
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8p23"
1..4580
/gene="defensin-related pseudogene"
/pseudo
BASE COUNT 1194 a 1175 c 950 g 1261 t
ORIGIN

Alignment Scores:
Pred. No.: 2 86e-05 Length: 4580
Score: 108.00 Matches: 24
Percent Similarity: 68.29% Conservative: 4
Best Local Similarity: 58.54% Mismatches: 13
Query Match: 46.35% Indels: 0
DB: 9 Gaps: 0

US-10-045-180A-5 (1-44) x AF195487 (1-4580)

Qy 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAsp 20
Db 3282 GAGCCACTCCAGGCAATGCTGATGAGGCTACAGCCAGGAGCAGCTGGAGCAGATGAT 3341

Qy 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 3342 CAGGAGTGGTTGATTCCTTTTCATGGGATGAAAGAGCTCTCTTCAGGTTTCAGGTAAG 3401

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 1538.36 Seconds
(without alignments)
695.153 Million cell updates/sec

Title: US-10-045-180A-5
Perfect score: 233
Sequence: 1 EPLQARAHMPAQKPPADD.....IYFSGDDSCSLQVPGSTKGL 44

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-Q=/cpu2 1/USPTO spool_p/US10045180/runat 17122003 145024 12540/app_query.fasta 1.860
-DB=EST -QFMT=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-LOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180 @CGN 1.1 6100 @runat 17122003 145024 12540 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -XGAPOP=10 -YGAPOP=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em estba:*
2: em esthum:*
3: em estin:*
4: em estmu:*
5: em estov:*
6: em estpl:*
7: em estro:*
8: em htc:*
9: gb est1:*
10: gb est2:*
11: gb htc:*
12: gb est3:*
13: gb est4:*
14: gb est5:*
15: em estfun:*
16: em estcom:*
17: em gss hum:*
18: em gss inv:*
19: em gss pln:*
20: em gss vrt:*
21: em gss fun:*
22: em gss mam:*
23: em gss mus:*
24: em gss pro:*
25: em gss rod:*
26: em gss phg:*
27: em gss vri:*
28: gb gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	114	48.9	307	12	BI021150 CM0-MT035
C 2	114	48.9	422	13	EX099382 BX099382
C 3	114	48.9	458	9	AI082774 ox7Aa06.x
C 4	114	48.9	541	9	AA722853 zg85h07.x
C 5	114	48.9	542	9	AI250799 qi36g07.x
C 6	114	48.9	634	13	BU616655 UI-H-DF0-
C 7	108	46.4	788	13	EX117347 BX117347
C 8	107	45.9	682	29	AG171073 Pan trogl
C 9	93.5	40.1	476	9	AI081728 ow81f07.s
C 10	81.5	35.0	319	12	BI021560 CM3-MT034
C 11	80	34.3	249	12	BM830493 K-BST0104
C 12	80	34.3	300	12	BM765919 K-BST0047
C 13	80	34.3	303	12	BM855729 K-BST0138
C 14	80	34.3	351	12	BM768678 K-BST0051
C 15	80	34.3	375	12	BM856063 K-BST0139
C 16	80	34.3	376	12	BM855276 K-BST0138
C 17	80	34.3	443	12	BM855593 K-BST0138
C 18	80	34.3	447	12	BM703419 UI-E-CL1-
C 19	80	34.3	460	12	BM768516 K-BST0051
C 20	80	34.3	467	12	BM854660 K-BST0137
C 21	80	34.3	511	12	BM715953 UI-E-EJ0-
C 22	80	34.3	513	12	BM674264 UI-E-EJ0-
C 23	80	34.3	519	13	BU729350 UI-E-CL1-
C 24	78.5	33.7	208	14	R29450 FI-57ID 22
C 25	78.5	33.7	281	9	AA321281 EST23811
C 26	78.5	33.7	294	9	AA321316 EST23783
C 27	78.5	33.7	306	10	BF993430 QV3-MT012
C 28	78.5	33.7	309	12	BI022856 CM4-MT024
C 29	78.5	33.7	312	9	AA321102 EST23896
C 30	78.5	33.7	320	9	AA321426 EST23923
C 31	78.5	33.7	325	10	BF993434 QV3-MT012
C 32	78.5	33.7	327	9	AA321192 EST23994
C 33	78.5	33.7	330	10	BE831394 QV4-MT005
C 34	78.5	33.7	348	10	BF993385 QV3-MT012
C 35	78.5	33.7	349	12	BI021617 CM3-MT034
C 36	78.5	33.7	354	9	AA321384 EST23855
C 37	78.5	33.7	355	10	BF993379 QV3-MT012
C 38	78.5	33.7	359	12	BI021220 CM0-MT035
C 39	78.5	33.7	370	10	BF993376 QV3-MT012
C 40	78.5	33.7	370	10	BF993383 QV3-MT012
C 41	78.5	33.7	371	10	BF993218 QV3-MT012
C 42	78.5	33.7	372	10	BF993378 QV3-MT012
C 43	78.5	33.7	377	9	AA321224 EST23726
C 44	78.5	33.7	379	12	BI021556 CM3-MT034
C 45	78.5	33.7	379	14	COL361 HUMGS000834

ALIGNMENTS

RESULT 1
BI021150/c
LOCUS CM0-MT0351-150101-736-b03 MT0351 Homo sapiens cDNA, mRNA sequence.
DEFINITION BI021150
ACCESSION BI021150
VERSION BI021150.1 GI:14427780
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 307)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,

TITLE
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE
20202663

PUBMED
10737800

COMMENT
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM0&tl2=CM0-MT0351-150101-736-b03&t3=2001-01-15&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 21
High quality sequence stop: 307.

FEATURES
Location/Qualifiers
1..307
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0351"
/note="Organ: marrow; Vector: puc18; Site:1; SmaI; Site:2; SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT
75 a 88 c 76 g 68 t

ORIGIN

Alignment Scores:
Pred. No.: 1.06e-05 Length: 307
Score: 114.00 Matches: 23
Percent Similarity: 69.77% Conservative: 7
Best Local Similarity: 53.49% Mismatches: 13
Query Match: 48.93% Indels: 0
DB: 12 Gaps: 0

US-10-045-180A-5 (1-44) x BI021150 (1-307)

Qy 2 ProLeuGlnAlaAArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspGln 21
|||||
Db 286 CCACCTCCAGGCAAGAGGTGATGAGCTCCAGCCAGGAGCAGCGTGGGCCAGAGACCAG 227
|||||

Qy 22 AspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySerThr 41
|||||
Db 226 GACATATCTATTTCCTTTGCATGGGATAAAGCTCTGCTCTTCAGTTTCAGGCTCAACA 167
|||||

Qy 42 LysGlyLeu 44
:::|
Db 166 AGGGGCATG 158
:::|

RESULT 2
BX099382 422 bp mRNA linear EST 06-FEB-2003
LOCUS BX099382
DEFINITION BX099382 Soares NhMPu_S1 Homo sapiens cDNA clone IMAGE998H134549 ;
IMAGE:1858620, mRNA sequence.
ACCESSION BX099382
VERSION BX099382.1 GI:27844054
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 422)

REFERENCE
AUTHORS
Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M., Radelof,U., Schneider,D. and Korn,B.
Human Unigeneset - RZPD3
Unpublished
Contact: Ina Rofls
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
RZPD: IMAGp998H134549
RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
RZPDLIB - RZPD3 (RZPDLIB No.972)
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rofls
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
Heubnerweg 6, D-14059 Berlin, Germany
Tel: +49 30 32639 101
Fax: +49 30 32639 111
www.rzpd.de
This clone is available royalty-free from RZPD;
contact RZPD (clone@rzpd.de) for further information. Seq primer:
M13r, Primer sequence: TTTCACACAGGAACAGCTATGAC.

FEATURES
Location/Qualifiers
1..422
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGp998H134549 ; IMAGE:1858620"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/clone_lib="Soares NhMPu_S1"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBHM, pregnant uterus NBHPU, and fetal heart NBHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT
83 a 115 c 115 g 107 t

ORIGIN

Alignment Scores:
Pred. No.: 1.58e-05 Length: 422
Score: 114.00 Matches: 23
Percent Similarity: 69.77% Conservative: 7
Best Local Similarity: 53.49% Mismatches: 13
Query Match: 48.93% Indels: 0
DB: 13 Gaps: 0

US-10-045-180A-5 (1-44) x BX099382 (1-422)

Qy 2 ProLeuGlnAlaAArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspGln 21
|||||
Db 130 CCACCTCCAGGCAAGAGGTGATGAGCTCCAGCCAGGAGCAGCGTGGGCCAGAGACCAG 189
|||||

Qy 22 AspValValIleTyrPheSerGlyAspSerCysSerLeuGlnValProGlySerThr 41
|||||
Db 190 GACATATCTATTTCCTTTGCATGGGATAAAGCTCTGCTCTTCAGTTTCAGGCTCAACA 249
|||||

Qy 42 LysGlyLeu 44
:::|
Db 250 AGGGGCATG 258
:::|

RESULT 3
BX099382 422 bp mRNA linear EST 24-SEP-1998
LOCUS BX099382
DEFINITION BX099382 Soares NhMPu_S1 Homo sapiens cDNA clone IMAGE998H134549 ;
IMAGE:1858620, mRNA sequence.
ACCESSION BX099382
VERSION BX099382.1 GI:27844054
KEYWORDS EST.
SOURCE Homo sapiens (human)

```

DEFINITION ox74a06.x1 Soares_NhMpu_S1 Homo sapiens cDNA clone IMAGE:1662034
3' similar to gb:X65977 NEUTROPHIL DEFENSIN 4 PRECURSOR (HUMAN);,
mRNA sequence.
ACCESSION AI082774
VERSION AI082774.1 GI:3417750
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1. (bases 1 to 458)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicapp.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 637 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 446.
FEATURES             Location/Qualifiers
     source            1..458
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1662034"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /clone_lib="Soares_NhMpu_S1"
                     /notes="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site 1: Not 1;
                     Site 2: Eco RI: Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NBHM, pregnant uterus
                     NBHM, and fetal heart NBHM9W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. Clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT   128 a   117 c   103 g   109 t   1 others
ORIGIN
Alignment Scores:
Pred. No.:      1.74e-05      Length:      458
Score:          114.00      Matches:      23
Percent Similarity: 69.77%      Conservative:  7
Best Local Similarity: 53.49%      Mismatches:   13
Query Match:      48.93%      Indels:       0
DB:              9          Gaps:         0
US-10-045-180A-5 (1-44) x AI082774 (1-458)
QY      2  ProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspGln 21
Db      431 CCACCTCCAGCAGAGGTGTATGAGCTCCAGGCCAGGAGCAGGTGGCCAGAGACCAG 372
QY      22 AspValVallIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 41
Db      371 GACATATCTATTTCCTTTGCATGGGATAAAAGCTCTGCTCTTCAGGTTTCAGGCTCAACA 312
QY      42  LysGlyLeu 44
Db      311 AGGGGCATG 303
RESULT 4
AA722853/c
LOCUS      AA722853
DEFINITION z98h07.s1 Soares fetal_heart NBHM19W Homo sapiens cDNA clone
IMAGE:409405 3' similar to gb:X65977 NEUTROPHIL DEFENSIN 4

```

```

RESULT 5
AI250799/c 542 bp mRNA linear EST 28-JAN-1999
LOCUS q136907.x1 Soares NHMPu S1 Homo sapiens cDNA clone IMAGE:1858620
DEFINITION 3' similar to gb:X65977 NEUTROPHIL DEFENSIN 4 PRECURSOR (HUMAN);,
          mRNA sequence.
ACCESSION AI250799
VERSION AI250799.1 GI:3847328
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLES 1 (bases 1 to 542)
NCl-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 633 Std Error: 0.00
Seq primer: -40bp from Gibco
High quality sequence stop: 463.
FEATURES
    source
        1..542
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:1858620"
            /tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
            /lab_host="DH10B"
            /clone_lib="Soares NHMPu S1"
            /note="Organ: mixed (see below); Vector: pRTT3D-Pac
(Pharmacia) with a modified polylinker; Site 1: Not 1;
Site 2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHMU, and fetal heart NBHM19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of 1 M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
BASE COUNT 151 a 134 c 136 g 119 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 2.15e-05 Length: 542
Score: 114.00 Matches: 23
Percent Similarity: 69.77% Conservative: 7
Best Local Similarity: 53.49% Mismatches: 13
Query Match: 48.93% Indels: 0
Gaps: 0
US-10-045-180a-5 (1-44) x AI250799 (1-542)
Cy 2 ProleuglnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAspGln 21
Db 431 CCACCTCCAGCAAGAGGTGATGAGCTCCAGGCGAGGAGCGTGGGCCGAGAAGACGAG 372
Cy 22 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 41
Db 371 GACATATCTATTTCCTTGATGAGGATAAAGAGCTCTGCTTCAGGTTTCAGGCTCAACA 312
Cy 42 LysGlyLeu 44
Db 311 AGGGGCAATG 303
RESULT 6
BU616655/c

```

```

LOCUS BU616655 634 bp mRNA linear EST 23-SEP-2002
DEFINITION UI-H-DFO-ben-n-14-0-UI-s1 NCl CGAP DFO Homo sapiens cDNA clone
UI-H-DFO-ben-n-14-0-UI 3', mRNA sequence.
ACCESSION BU616655
VERSION BU616655.1 GI:23282870
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLES 1 (bases 1 to 634)
NCl-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA-Yes.
FEATURES
    source
        1..634
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="UI-H-DFO-ben-n-14-0-UI"
            /tissue_type="Subchondral Bone"
            /dev stage="Adult"
            /lab_host="DH10B (Life Technologies)"
            /clone_lib="NCl CGAP DFO"
            /note="Organ: Bone; Vector: pRTT3-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
NCl CGAP DFO is a cDNA library containing the following
tissue(s): Subchondral Bone. The library was constructed
according to Bonaldo, Lemmon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pRTT3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dT)18 tail. The
sequence tag for this library is GTTAAAGCGTC.
TAG LIB=UI-H-DFO
TAG TISSUE=subchondral bone
TAG SEQ=GTAAAGCGTC"
BASE COUNT 170 a 151 c 158 g 155 t
ORIGIN
Alignment Scores:
Pred. No.: 2.62e-05 Length: 634
Score: 114.00 Matches: 23
Percent Similarity: 69.77% Conservative: 7
Best Local Similarity: 53.49% Mismatches: 13
Query Match: 48.93% Indels: 0
Gaps: 0
US-10-045-180a-5 (1-44) x BU616655 (1-634)
Cy 2 ProleuglnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAspGln 21
Db 449 CCACCTCCAGCAAGAGGTGATGAGGCTCCAGGCGAGGAGCGTGGGCCGAGAAGACGAG 390
Cy 22 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 41
Db 389 GACATATCTATTTCCTTGATGAGGATAAAGAGCTCTGCTTCAGGTTTCAGGCTCAACA 330
Cy 42 LysGlyLeu 44

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 165 Seconds
(without alignments)
719.850 Million cell updates/sec

Title: US-10-045-180a-5
Perfect score: 233
Sequence: 1 EPIQARAHMPAQKPPADD.....IYSGDDSCSLQVPSSTKGL 44

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2552756 segs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-O=/cgr2.1/usfto.spool.p/us10045180/runat.17122003.145023.12520/app.query.fasta_1.860
-DB=N Geneseq.19jun03 -QFMT=fastap -SUFFIX=mg -MINMATCH=0.1 -LOOPT=0
-LOOPT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human0.cdi
-LIST=45 -DOCLINK=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=45
-MODE=LOCAL -OUTFMT=pct -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180 @CN 1.1 835 @runat.17122003.145023.12520 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -MAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.19jun03.*

1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	233	100.0	453	20	AA26697
2	233	100.0	453	20	AA26698
3	209	89.7	4415	20	AA26696
4	114	48.9	860	25	ABX63473
5	104	44.6	496	25	AA251767
6	101	43.3	294	22	AA57220
7	101	43.3	452	14	AA053217
8	101	43.3	452	25	ACAS6600
9	90	38.6	398	24	ABNS9997
10	86	36.9	243	22	AA666930
11	83	35.6	495	22	AA666925
12	83	35.6	495	25	AA251769
13	82	35.2	500	22	AA666924
14	82	35.2	500	25	AA251768
15	80	34.3	424	14	AA053216
16	80	34.3	449	24	AB060783
17	80	34.3	514	22	AA57427
18	80	34.3	586	24	AB060782
19	80	34.3	2548	22	AA666927
20	78.5	33.7	299	16	AA726106
21	78.5	33.7	464	21	AA20911
22	78.5	33.7	464	21	AA34789
23	78.5	33.7	478	21	AA04468
24	78.5	33.7	498	12	AA014808
25	78.5	33.7	498	24	AB080601
26	78.5	33.7	514	24	ABNS9722
27	78.5	33.7	543	25	AB236767
28	78.5	33.7	644	21	AA21446
29	78.5	33.7	652	24	AB211202
30	78.5	33.7	654	24	AB21183
31	78.5	33.7	209273	21	AA21437
32	76	32.6	2523	22	AA666926
33	76	32.6	2880	14	AA053218
34	74	31.8	445	17	AA730734
35	74	31.8	2408	17	AA730737
36	73.5	31.5	473	25	AB236740
37	73	31.3	411	24	AB060009
38	72.5	31.1	652	24	AB211202
39	72	30.9	2457	17	AA730736
40	72	30.9	2551	17	AA730738
41	71	30.5	445	17	AA730733
42	71	30.5	445	17	AA730735
43	70	30.0	480	25	AA247375
44	69	29.6	412	24	AB239009
45	69	29.6	412	25	AC11338

ALIGNMENTS

RESULT 1
AA26697
ID AAX26697 standard; cDNA; 453 BP.

AC AAX26697;

DT 18-JUN-1999 (first entry)

DB cDNA sequence encoding human defensin (Def-X) protein.

KX Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;

KW cytotoxic; anticancer; inflammation; tissue repair;

KM endocrine regulation; corticosteroid regulation; cancer; melanoma;

OS AIDS; immune deficiency; psoriasis; ss.

XX Homo sapiens.

PN FR2767832-A1.
XX 05-MAR-1999.
XX 29-AUG-1997; 97FR-0010823.
XX 29-AUG-1997; 97FR-0010823.
XX (GEST) GENSET SA.
XX Bougueleret I, Chumakov I;
XX WPI; 1999-183266/16.
XX P-PSDB; AAY01604.
PT Human defensin polypeptide Def-X - useful as antimicrobial agent,
PT anticancer agent, pesticide, etc.
PS Disclosure; Fig 3; 56pp; French.
XX The present sequence encodes human defensin (Def-X). The Def-X
XX polypeptide can be used as an antimicrobial, antiparasitic agent or
XX a pesticide. The Def-X polypeptide can be used as a cytostatic
XX (especially anticancer) agent, and as an agent for modulating processes
XX of inflammation, tissue repair and endocrine (especially corticosteroid)
XX regulation. The polypeptide can be used in a composition for external
XX topical use, especially in a cosmetic composition. Compositions
XX containing Def-X can be used for prevention and treatment of microbial
XX and parasitic infections, especially where the microbial or parasitic
XX infections are Gram-positive or -negative bacterial infections or
XX mycobacterial, fungal or spirochaete infections, or where the viral
XX infections are associated with enveloped viruses, especially HSV and HIV.
XX The compositions can be used for prevention and/or treatment of cancers,
XX especially melanomas, or liver cancer, prostate cancer, non-small-cell
XX lung cancer or colorectal carcinoma, and for enhancing immunity,
XX especially in the case of AIDS, or preventing immune deficiency,
XX inflammatory processes, especially in the case of chronic inflammatory
XX disorders.
SQ Sequence 453 BP; 107 A; 121 C; 97 G; 128 T; 0 other;
Alignment Scores:
Pred. No.: 2.62e-25 Length: 453
Score: 233.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-10-045-180A-5 (1-44) x AAX26697 (1-453)
QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnIysGlnProProAlaAspAsp 20
Db 109 GAGCCGCTCCAGGCAAGAGCTCATGATGCCAGCCCAAGACGCTCCAGAGATGAC 168
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 169 CAGGATGTGCTATTACTTTTCAGAGATGACAGCTGCTCTTTCAGGTTCCAGGCTCA 228
QY 41 ThrIysGlyLeu 44
Db 229 ACAAAGGCTTG 240
RESULT 2
AAX26698
ID AAX26698 standard; cDNA; 453 BP.
XX AAX26698;
AC AAX26698;
XX 18-JUN-1999 (first entry)
DT 18-JUN-1999 (first entry)
XX cDNA sequence encoding human defensin (Def-X) protein.

XX Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
XX cytostatic; anticancer; inflammation; tissue repair;
XX endocrine regulation; corticosteroid regulation; cancer; melanoma;
XX AIDS; immune deficiency; psoriasis; ss.
OS Homo sapiens.
XX FR2767832-A1.
XX 05-MAR-1999.
XX 29-AUG-1997; 97FR-0010823.
XX 29-AUG-1997; 97FR-0010823.
XX 29-AUG-1997; 97FR-0010823.
XX (GEST) GENSET SA.
XX Bougueleret I, Chumakov I;
XX WPI; 1999-183266/16.
DR Human defensin polypeptide Def-X - useful as antimicrobial agent,
PT anticancer agent, pesticide, etc.
PS Claim 5; Fig 4; 56pp; French.
XX The present sequence encodes human defensin (Def-X). The Def-X
XX polypeptide can be used as an antimicrobial, antiparasitic agent or
XX a pesticide. The Def-X polypeptide can be used as a cytostatic
XX (especially anticancer) agent, and as an agent for modulating processes
XX of inflammation, tissue repair and endocrine (especially corticosteroid)
XX regulation. The polypeptide can be used in a composition for external
XX topical use, especially in a cosmetic composition. Compositions
XX containing Def-X can be used for prevention and treatment of microbial
XX and parasitic infections, especially where the microbial or parasitic
XX infections are Gram-positive or -negative bacterial infections or
XX mycobacterial, fungal or spirochaete infections, or where the viral
XX infections are associated with enveloped viruses, especially HSV and HIV.
XX The compositions can be used for prevention and/or treatment of cancers,
XX especially melanomas, or liver cancer, prostate cancer, non-small-cell
XX lung cancer or colorectal carcinoma, and for enhancing immunity,
XX especially in the case of AIDS, or preventing immune deficiency,
XX inflammatory processes, especially in the case of chronic inflammatory
XX disorders.
SQ Sequence 453 BP; 107 A; 121 C; 97 G; 128 T; 0 other;
Alignment Scores:
Pred. No.: 2.62e-25 Length: 453
Score: 233.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
US-10-045-180A-5 (1-44) x AAX26698 (1-453)
QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnIysGlnProProAlaAspAsp 20
Db 109 GAGCCGCTCCAGGCAAGAGCTCATGATGCCAGCCCAAGACGCTCCAGAGATGAC 168
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 169 CAGGATGTGCTATTACTTTTCAGAGATGACAGCTGCTCTTTCAGGTTCCAGGCTCA 228
QY 41 ThrIysGlyLeu 44
Db 229 ACAAAGGCTTG 240
RESULT 3
AAX26696

ID AAX26696 standard; DNA; 4415 BP.
 XX AAX26696;
 AC
 XX 18-JUN-1999 (first entry)
 DT
 XX
 XX Nucleotide sequence of human defensin (Def-X).
 DE
 XX Human defensin; Def-X; antimicrobial; antiparasitic; peptidic;
 KW cytostatic; anticancer; inflammation; tissue repair;
 KW endocrine regulation; corticostatic regulation; cancer; melanoma;
 KW AIDS; immune deficiency; psoriasis; ss.
 KM
 OS Homo sapiens.
 XX
 XX FR2767832-A1.
 PN
 XX 05-MAR-1999.
 PD
 XX 29-AUG-1997; 97FR-0010823.
 PF
 XX 29-AUG-1997; 97FR-0010823.
 PR
 XX 29-AUG-1997; 97FR-0010823.
 XX
 XX (GEST) GENSET SA.
 PA
 XX Bougueleret L, Chumakov I;
 PI
 XX WPI; 1999-183266/16.
 DR
 XX Human defensin polypeptide Def-X - useful as antimicrobial agent,
 PT anticancer agent, pesticide, etc.
 PT
 XX Claim 5; Fig 2; 56pp; French.
 PS
 XX The present sequence encodes human defensin (Def-X). The Def-X
 CC polypeptide can be used as an antimicrobial, antiparasitic agent or
 CC a pesticide. The Def-X polypeptide can be used as a cytostatic
 CC (especially anticancer) agent, and as an agent for modulating processes
 CC of inflammation, tissue repair and endocrine (especially corticostatic)
 CC regulation. The polypeptide can be used in a composition for external
 CC topical use, especially in a cosmetic composition. Compositions
 CC containing Def-X can be used for prevention and treatment of microbial
 CC and parasitic infections, especially where the microbial or parasitic
 CC infections are Gram-positive or -negative bacterial infections or
 CC mycobacterial, fungal or spirochaete infections, or where the viral
 CC infections are associated with enveloped viruses, especially HSV and HIV.
 CC The compositions can be used for prevention and/or treatment of cancers,
 CC especially melanomas, or liver cancer, prostate cancer, non-small-cell
 CC lung cancer or colorectal carcinoma, and for enhancing immunity,
 CC especially in the case of AIDS, or preventing immune deficiency,
 CC inflammatory processes, especially in the case of chronic inflammatory
 CC disorders.
 CC
 SQ Sequence 4415 BP; 1128 A; 1109 C; 912 G; 1252 T; 14 other;
 Alignment Scores:
 Pred. No.: Length: 4415
 Score: 209.00 Matches: 39
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 89.70% Indels: 0
 DB: 20 Gaps: 0
 US-10-045-180A-5 (1-44) x AAX26696 (1-4415)
 QY 1 GIUProLeuGlnAlaIrgAlaHisGluMetProAlaGlnIlyGlnProAlaAspAsp 20
 DB 3463 GAGCCGCTCCAGGCAAGAGCTCATGAGATGCCAGCCAGAGAGAGCTCCAGAGATGAC 3522
 QY 21 GlnAspValValIleIyrheserGlyAspAspSerCysSerLeuGlnValProGly 39
 DB 3523 CAGGATGTGTCTACTTTCAGAGATGACAGCTGCTCTTCAGGTTCCAGGT 3579

RESULT 4
 ID ABX63473 standard; cDNA; 860 BP.
 XX
 XX ABX63473;
 AC
 XX 26-FEB-2003 (first entry)
 DT
 XX Human cDNA #473 differentially expressed in activated vascular tissue.
 DE
 XX
 XX Human; gene; ss; vascular tissue; cytostatic; atherosclerosis;
 KM cardiant; hypotensive; antidiabetic; gynaecological; vasotropic;
 KM cerebroprotective; gene therapy; vascular disease; cancer; coronary;
 KM artery disease; hypertension; diabetes; pre-eclampsia; restenosis;
 KM ischaemia-reperfusion injury; stroke;
 KM
 XX
 XX Homo sapiens.
 OS
 XX US2002137081-A1.
 PN
 XX 26-SEP-2002.
 PD
 XX 08-JAN-2002; 2002US-0044090.
 PF
 XX 28-JUL-2000; 2000US-222469P.
 PR
 XX 08-JAN-2001; 2001US-260483P.
 XX
 XX (BAND/) BANDMAN O.
 PA
 XX Bandman O;
 PI
 XX WPI; 2003-110597/10.
 DR
 XX Combination for diagnosing, staging, treating, or monitoring the
 PT progression of treatment of a vascular disease, e.g. atherosclerosis.
 PT comprises several cDNAs that are differentially expressed in activated
 PT vascular tissue -
 PS
 XX Claim 1; Page -; 18pp; English.
 XX This invention relates to a combination comprising several cDNAs that
 CC are differentially expressed in activated vascular tissue. The invention
 CC also discloses a high throughput method for detecting differentially
 CC expressed cDNAs in a sample. The cDNAs of the invention may have
 CC antiarteriosclerotic, cytostatic, cardiant, hypotensive, antidiabetic;
 CC gynaecological, vasotropic and cerebroprotective activities and may be
 CC used in gene therapy. The cDNAs of the invention may be used in a
 CC high-throughput methods for detecting differential expression of one or
 CC more cDNAs in a sample, or screening several molecules or compounds to
 CC identify a molecule or compound that specifically binds a cDNA of the
 CC invention. A protein encoded by the cDNA may be used to screen several
 CC molecules or compounds to identify a ligand that specifically binds to
 CC the protein, or to produce or purify an antibody to the protein that can
 CC be used to detect a protein in a sample or purify a natural or
 CC recombinant protein from a sample. The nucleotides may be useful for
 CC diagnosing, staging, treating, or monitoring the progression of
 CC treatment of a vascular disease, e.g. atherosclerosis, cancer, coronary
 CC artery disease, hypertension, diabetes, pre-eclampsia, ischaemia-
 CC reperfusion injury, restenosis, or stroke. The cDNAs can also be used
 CC for large-scale genetic or gene expression analysis of several new
 CC nucleic acid molecules. Antibodies to the proteins encoded by the
 CC cDNAs are useful for diagnosing pre-pathologic disorders, and chronic
 CC or acute diseases associated with abnormalities in the expression,
 CC amount or distribution of the protein. The present sequence
 CC represents a cDNA of the invention that is differentially expressed in
 CC activated vascular tissue.
 CC Note: The sequence data for this patent did not form part of the
 CC specification, but was obtained in electronic format directly from USPTO
 CC at <http://seqdata.uspto.gov/sequence.html?docID=20020137081>.
 XX
 SQ Sequence 860 BP; 216 A; 211 C; 192 G; 241 T; 0 other;

Alignment Scores:

Pred. No.: 2,78e-07 Length: 860
 Score: 114.00 Matches: 23
 Percent Similarity: 69.77% Conservative: 7
 Best Local Similarity: 53.49% Mismatches: 13
 Query Match: 48.93% Indels: 0
 DB: 25 Gaps: 0

US-10-045-180a-5 (1-44) x ABX63473 (1-860)

QY 2 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlysgInProProAlaAspAspGln 21
 DB 136 CCACGCCAGGAGGAGGATGAGGCTCCAGGCCAGAGACACCGTGGCCAGAACACCG 195
 QY 22 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 41
 DB 196 GACATATCTATTTCTTTGATGAGATGAAGCTGCTTCAGGCTTCAGGCTCAACA 255

QY 42 LysGlyLeu 44
 DB 256 AGGGGATG 264

RESULT 5
 ID AAD51767 standard; DNA; 496 BP.
 XX

AC AAD51767;
 DT 16-APR-2003 (first entry)
 DE Human retrocyclin DNA.
 XX

KW Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
 KW human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
 KW antibiotic modelling; antimicrobial; human; gene; ds.
 XX

OS Homo sapiens.

FH Key Location/Qualifiers
 FT CDS 63..495

FT /*tag= a
 FT /product= "Human retrocyclin precursor protein"
 FT /transl_except= (pos:124..126, aa:Xaa)
 FT /transl_except= (pos:304..306, aa:Xaa)
 FT /transl_except= (pos:361..363, aa:Xaa)
 FT /transl_except= (pos:409..411, aa:Xaa)
 FT /note= "Xaa corresponds to in-frame stop codon;
 FT No start and stop codon"
 FT sig_peptide 63..126
 FT /*tag= b
 FT mat_peptide 127..495
 FT /*tag= c
 FT /product= "Mature human retrocyclin protein"
 FT CDS 127..306
 FT /*tag= d
 FT /product= "Human mature retrocyclin protein fragment"
 XX

PN W0200285401-A1.

XX 31-OCT-2002.

XX 18-APR-2002; 2002WO-US12353.

XX 18-APR-2001; 2001US-284855P.

XX (REGC) UNIV CALIFORNIA.

XX PI Lehrer RI, Waring AJ, Cole AM, Hong TB;

XX WPI; 2003-103387/09.

XX P-PSDB; AAE33811, AAE33860.

PT New isolated retrocyclin peptide, useful for preventing retroviral
 PT infections in cells susceptible to bacterial or viral infections or
 PT treating patients having the infections, such as HIV, sexually
 PT transmitted diseases, vaginosis
 XX Claim 14; Page 50-51; 72pp; English.

CC The invention relates to novel retrocyclin peptides. Peptides and
 CC methods of the invention are useful for preventing retroviral infections
 CC in cells susceptible to bacterial or viral infections, or treating
 CC patients having infections such as HIV (human immunodeficiency virus),
 CC sexually transmitted diseases, bacterial vaginosis or ophthalmic
 CC infections. The retrocyclin-mediated killing is useful for modeling
 CC and screening novel antibiotics. The invention is also useful in gene
 CC therapy. The present sequence is human retrocyclin DNA.
 XX

SQ Sequence 496 BP; 106 A; 138 C; 117 G; 135 T; 0 other;

Alignment Scores:

Pred. No.: 4.22e-06 Length: 496
 Score: 104.00 Matches: 22
 Percent Similarity: 68.18% Conservative: 8
 Best Local Similarity: 50.00% Mismatches: 14
 Query Match: 44.64% Indels: 0
 DB: 25 Gaps: 0

US-10-045-180a-5 (1-44) x AAD51767 (1-496)

QY 1 GlnProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlysgInProProAlaAspAsp 20

DB 136 GAGCCACTTCAGGAGGAGGATGAGGCTCCAGGCCAGAGACACCGTGGCCAGAACACCG 195

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
 DB 196 CAGGAATGCTCATGCTTATCATGAGCATGAAGTCCGCTTCGCTTCAGACTCA 255

QY 41 ThrLysGlyLeu 44

DB 256 GCGAGAGGCTTG 267

RESULT 6

ID AAH57220 standard; cDNA; 294 BP.

XX AAH57220;

DT 10-SEP-2001 (first entry)

XX Human intestine specific cDNA sequence SEQ ID NO:60.

KW Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
 KW lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; es;
 KW metabolic disease; developmental disease; cytostatic; immunomodulatory;
 KW neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
 XX

OS Homo sapiens.

PN W0200132927-A2.

XX 10-MAY-2001.

XX 02-NOV-2000; 2000WO-US30396.

XX 04-NOV-1999; 99US-0163508.

XX (INCY-) INCYTE GENOMICS INC.

XX Sornase T, Sellhamey JU, Watson GA;

XX WPI; 2001-291057/30.

PT New cell and tissue specific polynucleotides useful for diagnosis,
 PT prognosis or monitoring of treatments for disorders where the gene is

PT associated with a cancer, immunopathology or neuropathology -
XX
PS Claim 1, Page 95; 327pp; English.
XX
CC AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC sequences (I). (I) can have cytostatic, immunomodulatory and
CC neuroprotective activities, and can be used in gene therapy. (I) and
CC proteins (II) encoded by then are used in high throughput screening
CC assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC their fragments, immunoglobulins, inhibitors, drug compounds and
CC pharmaceutical agents. Expression of (I) in a sample indicates the
CC differentiation of embryonic stem cells into a tissue selected from
CC brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC tissues. (I) and (II) are used to produce an expression profile that
CC defines a metabolic or developmental process, treatment, condition,
CC disease or disorder. The gene profile can be used for diagnosis,
CC prognosis or monitoring of treatments and for investigating a
CC predisposition to a disorder where the gene is associated with a
CC cancer, immunopathology or neuropathology.
XX
SQ Sequence 294 BP; 72 A; 80 C; 73 G; 64 T; 5 other;

Alignment Scores:
Pred. No.: 6.04e-06 Length: 294
Score: 101.00 Matches: 21
Percent Similarity: 71.43% Conservative: 9
Best Local Similarity: 50.00% Mismatches: 12
Query Match: 43.35% Indels: 0
DB: Gaps: 0

US-10-045-180a-5 (1-44) x AAH57220 (1-294)

QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnIysGlnProProAlaAspAsp 20
Db 93 GATCCACTGCGAGGAAAGCTTATGAGGCTGATGCCAGAGCAAGCGTGGGCAATGAC 152

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 153 CAGGACTTGGCCGTCCTTTCCTTGCAGAGGATCAAGCTCAAGCTTTCAGCTTGGGCTCA 212

QY 41 ThrIys 42
Db 213 ACAAGG 218

RESULT 7
AAQ53217
ID AAQ53217 standard; cDNA; 452 BP.
XX
AC AAQ53217;
XX
DT 25-MAR-2003 (updated)
DT 17-JUN-1994 (first entry)
XX
DE Sequence encoding gastrointestinal defensin (GID) peptide called
DE human defensin 6.
XX
KW Gastrointestinal defensin peptide; GID; pharmaceutical; Paneth
KW cell; antimicrobial; anti-inflammatory; diagnosis;
KM contact disinfectant; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 19..321
FT /tag= a
XX
PN W09324513-A1.
XX
PD 09-DEC-1993.
XX
PE 18-MAY-1993; 93WO-US04740.
XX

PR 22-MAY-1992; 92US-0888232.
XX
XX (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
XX
PI Bevins CL, Jones DE;
XX
DR WPI; 1993-405719/50.
DR P-PsDB; AAR44819.
XX
XX Gastrointestinal defensin peptide(s) - useful as antimicrobial
XX and anti-inflammatory agents and for detecting gastrointestinal
XX disorders
PS Claim 4; page 48-49; 97pp; English.
XX
XX A probe (D5' oligo) having the sequence given in AAQ53219 (claimed)
XX was used to screen a population of clones to identify novel defensin
XX defensin peptides. Pref. the libraries were human genomic and cDNA
XX libraries. Hybridisation and partial sequence analysis of the
XX identified clones contained previously characterised myeloid derived
XX defensin sequences as well as new defensin related sequences. Two
XX clones expressing new defensin related sequences were extensively
XX characterised and found to contain genes selectively expressed in
XX Paneth cells of the small intestine. These Paneth cell-derived
XX defensins are designated human defensin 5 and human defensin 6 and
XX are referred to as gastrointestinal defensin peptides. Comparison
XX of the deduced AA sequences of the defensin 5 cDNA with the
XX previously reported preprodefensin shows significant similarity.
XX The deduced AA sequence of defensin 6 cDNA has features similar to
XX defensin 5 and the previously reported preprodefensins.
XX (Updated on 25-MAR-2003 to correct PN field.)
SQ
Sequence 452 BP; 123 A; 118 C; 98 G; 113 T; 0 other;

Alignment Scores:
Pred. No.: 1.04e-05 Length: 452
Score: 101.00 Matches: 21
Percent Similarity: 71.43% Conservative: 9
Best Local Similarity: 50.00% Mismatches: 12
Query Match: 43.35% Indels: 0
DB: Gaps: 0

US-10-045-180a-5 (1-44) x AAQ53217 (1-452)

QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnIysGlnProProAlaAspAsp 20
Db 97 GATCCACTGCGAGGAAAGCTTATGAGGCTGATGCCAGAGCAAGCGTGGGCAATGAC 156

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 157 CAGGACTTGGCCGTCCTTTCCTTGCAGAGGATCAAGCTCAAGCTTTCAGCTTGGGCTCA 216

QY 41 ThrIys 42
Db 217 ACAAGG 222

RESULT 8
ACA56600
ID ACA56600 standard; cDNA; 452 BP.
XX
AC ACA56600;
XX
DT 06-JUN-2003 (first entry)
DT
XX
DE Human signalling pathway polynucleotide probe SEQ ID NO 1198.
XX
XX Human; probe; ss; array element; Parkinson's disease;
KW signalling pathway population; cancer; adenocarcinoma; leukaemia;
KM immunopathy; AIDS; asthma; neuropathy; Alzheimer's disease; microarray.
XX
OS Homo sapiens.
XX
XX US6500938-B1.
PN

```
XX
PD 31-DEC-2002.
XX
PF 30-JAN-1998; 98US-0016434.
XX
PR 30-JAN-1998; 98US-0016434.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI Au-Young J, Seilhamer JJ;
XX
DR WPI; 2003-352189/33.
XX
PT Combination of polynucleotide probes, useful as array elements in a
PT microarray for monitoring the expression of a number of target
PT polynucleotides
XX
PS Claim 1; SEQ ID NO 1198; 65pp; English.
XX
CC The invention relates to a combination which, comprises a number of
CC polynucleotide probes comprising a sequence selected from one of the 1490
CC sequences mentioned in the specification. The combination is useful as an
CC array element in a microarray for monitoring the expression of a number
CC of target polynucleotides. The microarray is particularly useful in the
CC diagnosis and treatment of cancer and immunopathology and neuropathology.
CC The microarray is useful in diagnostics and treatment regimens, drug
CC discovery and development, toxicological and carcinogenicity studies,
CC forensics and pharmacogenomics. The microarray is also useful for
CC monitoring progression of diseases and for developing sophisticated
CC profiles for the effects of currently available therapeutic drugs. The
CC combination is also useful for purifying a subpopulation of mRNAs, cDNAs
CC and genomic fragments and in research and diagnostic applications. The
CC array can detect changes in expression in a large number of genes coding
CC for different signaling pathway populations which can be used to diagnose
CC various diseases including cancer e.g. adenocarcinoma and leukaemia,
CC immunopathies e.g. AIDS and asthma, neuropathies e.g. Alzheimer's disease
CC and Parkinson's disease. The present sequence represents a polynucleotide
CC probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docid=06500938B1.
XX
SQ Sequence 452 BP; 124 A; 118 C; 98 G; 112 T; 0 other;

Alignment Scores:
Pred. No.: 1.04e-05 Length: 452
Score: 101.00 Matches: 21
Percent Similarity: 71.43% Conservative: 9
Best Local Similarity: 50.00% Mismatches: 12
Query Match: 43.35% Indels: 0
DB: 25 Gaps: 0

US-10-045-180A-5 (1-44) x ACA56600 (1-452)
QY 1 GlnProLeuGlnAlaArgAlaHisGluMetProAlaGlnIyGlnProProAlaAspAsp 20
Db 97 GATCCACTGAGGCAAAAGCTTATGAGGCTGATGCCAGAGCAAGGCTGGGCAAAATGAC 156
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 157 CAGGACTTGGCCGTCTCTTCCTTTCAGAGAGATGCAAGCTCAAGCTTATGAGCTTTGGGCTCA 216
QY 41 ThrLys 42
Db 217 ACAAGG 222

RESULT 9
ID AEN59997
AC AEN59997 standard; cDNA, 398 BP.
XX
XX 28-JUN-2002 (first entry)
DT
```

```
XX
DE Novel human coding sequence SEQ ID NO: 408.
XX
XX Human; antianaemic; vulnery; antiinflammatory; immunomodulator;
XX antileptility; cerebroprotective; cytostatic; rheumatic; gene therapy;
XX neuroprotective; antiparkinsonian; protein therapy; EST;
XX expressed sequence tag; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200222660-A2.
XX
PD 21-MAR-2002.
XX
PF 10-SEP-2001; 2001WO-US26015.
XX
PR 11-SEP-2000; 2000US-0659671.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F,
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;
XX
DR WPI; 2002-292408/33.
DR P-PDB; ABB97584.
XX
XX An isolated polynucleotide for treating diseases associated with its
XX encoded polypeptide such as cancer and multiple sclerosis -
XX
PS Claim 1; SEQ ID NO 408; 509pp; English.
XX
CC The present invention provides the protein and coding sequences of 444
CC novel human proteins. These were isolated from expressed sequences tags
CC (ESTs). They can be used to stimulate cell growth, to regulate
CC haematopoiesis e.g. to treat aplastic anemia, to help tissue regrowth
CC e.g. in burn treatment, to regulate the immune system e.g. to treat
CC multiple sclerosis, to regulate activin or inhibin e.g. to treat
CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
CC Parkinson's disease. The present sequence is a coding sequence of the
CC invention.
XX
SQ Sequence 398 BP; 102 A; 108 C; 93 G; 95 T; 0 other;

Alignment Scores:
Pred. No.: 0.000382 Length: 398
Score: 90.00 Matches: 19
Percent Similarity: 69.23% Conservative: 8
Best Local Similarity: 48.72% Mismatches: 12
Query Match: 38.63% Indels: 0
DB: 24 Gaps: 0

US-10-045-180A-5 (1-44) x AEN59997 (1-398)
QY 1 GlnProLeuGlnAlaArgAlaHisGluMetProAlaGlnIyGlnProProAlaAspAsp 20
Db 108 GATCCACTGAGGCAAAAGCTTATGAGGCTGATGCCAGAGCAAGGCTGGGCAAAATGAC 167
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 168 CAGGACTTGGCCGTCTCTTCCTTTCAGAGAGATGCAAGCTCAAGCTTATGAGCTTTGGGT 224

RESULT 10
ID AAC66930
AC AAC66930;
XX
XX 27-MAR-2001 (first entry)
DT
XX
XX Human theta defensin coding sequence SEQ ID NO: 28.
DT
```

```

KW Theta defensin; antimicrobial; cyclic; bacterium; fungus; protozoan;
KM virus; helminth; disinfectant; food preservative; analogue; ss.
XX
XX Homo sapiens.
XX
XX WO200068265-A1.
XX
XX 16-NOV-2000.
XX
XX 10-MAY-2000; 2000WO-US12842.
XX
XX 10-MAY-1999; 99US-0309487.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Selected ME, Tang Y, Yuan J, Ouellette AJ;
XX
XX WPI; 2001-031853/04.
XX
XX Novel theta defensin peptide with antimicrobial activity against
XX bacteria, yeast, fungi, protozoa and viruses
XX
XX Example 5; Fig 15; 110pp; English.
XX
XX The present invention provides theta defensin peptides and analogues
XX which have antimicrobial activity. They can be used in the treatment of
XX bacterial, viral, fungal, protozoan and helminthic infections, in
XX disinfectants and as food preservatives.
XX
XX Sequence 243 BP; 52 A; 75 C; 65 G; 51 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.0008 Length: 243
Score: 86.00 Matches: 20
Percent Similarity: 61.36% Conservative: 7
Best Local Similarity: 45.45% Mismatches: 17
Query Match: 36.91% Indels: 0
DB: 22 Gaps: 0
US-10-045-180A-5 (1-44) x AAC66925 (1-243)
QY 1 GIUPROLEUGINALARGALAHISGLUMETPROALAGINLYGInPROBROALASPAP 20
DB 67 GAGGACGCTCAGGCAAGAGCTGATGAGCGCCGACGACAGCCTCGAGCAATGAT 126
QY 21 GINASPVALIALLTYRPHESERGLYASPPASERCYSERLEUGINVALPROGLYSER 40
DB 127 CAGGGAATGCTCATTCCTTTACATGAGCCTGAAGACGCCGCTCTTCCACTTCAAGATCA 186
QY 41 ThrlysglyLeu 44
DB 187 GCGAAGGCTTG 198
RESULT 11
AAC66925
ID AAC66925 standard; cDNA; 495 BP.
XX
XX AAC66925;
XX
XX 27-MAR-2001 (first entry)
XX
XX Rhesus macaque theta defensin RTD-1b coding sequence.
XX
XX Theta defensin; antimicrobial; cyclic; bacterium; fungus; protozoan;
XX virus; helminth; disinfectant; food preservative; analogue; ss.
XX
XX Rhesus macaque.
XX
XX WO200068265-A1.
XX
XX 16-NOV-2000.
XX
XX 10-MAY-2000; 2000WO-US12842.

```

```

XX
XX 10-MAY-1999; 99US-0309487.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Selected ME, Tang Y, Yuan J, Ouellette AJ;
XX
XX WPI; 2001-031853/04.
XX
XX P-PSDB; AAB35041.
XX
XX Novel theta defensin peptide with antimicrobial activity against
XX bacteria, yeast, fungi, protozoa and viruses
XX
XX Example 5; Fig 11; 110pp; English.
XX
XX The present invention provides theta defensin peptides and analogues
XX which have antimicrobial activity. They can be used in the treatment of
XX bacterial, viral, fungal, protozoan and helminthic infections, in
XX disinfectants and as food preservatives.
XX
XX Sequence 495 BP; 103 A; 145 C; 122 G; 125 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.00553 Length: 495
Score: 83.00 Matches: 19
Percent Similarity: 61.36% Conservative: 8
Best Local Similarity: 43.18% Mismatches: 17
Query Match: 35.62% Indels: 0
DB: 22 Gaps: 0
US-10-045-180A-5 (1-44) x AAC66925 (1-495)
QY 1 GIUPROLEUGINALARGALAHISGLUMETPROALAGINLYGInPROBROALASPAP 20
DB 150 GAGGACGCTCAGGCAAGAGCTGATGAGCGCCGACGACAGCCTCGAGCAATGAT 209
QY 21 GINASPVALIALLTYRPHESERGLYASPPASERCYSERLEUGINVALPROGLYSER 40
DB 210 CAGGGAATGCTCATTCCTTTACATGAGCCTGAAGACGCCGCTCTTCCGCTTCAAGATCA 269
QY 41 ThrlysglyLeu 44
DB 270 GCGAAGGCTTG 281
RESULT 12
AADS1769
ID AADS1769 standard; DNA; 495 BP.
XX
XX AADS1769;
XX
XX 16-APR-2003 (first entry)
XX
XX Macaca mulatta RTD2 DNA.
XX
XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
XX human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
XX antibiotic modelling; antimicrobial; rhesus monkey; theta defensin 1b;
XX RTD2; gene; ds.
XX
XX Macaca mulatta.
XX
XX Key Location/Qualifiers
XX CDS 90..320
XX sig_peptide /product= "RTD2 protein"
XX mat_peptide /tag= b
XX /tag= C
XX /product= "Mature RTD2 protein"
XX
XX WO200285401-A1.

```

PD	31-OCT-2002.				
XX					
PF	18-APR-2002; 2002MO-US12353.				
XX					
PR	18-APR-2001; 2001US-284855P.				
XX					
PA	(REGC) UNIV CALIFORNIA.				
XX					
PI	Lehrer R.I, Waring A.J, Cole A.M, Hong T.B,				
XX					
DR	WPI; 2003-103387/09.				
XX	P-PSDB; AAE33813.				
PT	New isolated retrocyclin peptide, useful for preventing retroviral				
XX	infections in cells susceptible to bacterial or viral infections or				
PT	treating patients having the infections, such as HIV, sexually				
XX	transmitted diseases, vaginosis				
PS	-				
XX	Disclosure; Page 53; 72pp; English.				
CC	The invention relates to novel retrocyclin peptides. Peptides and				
CC	methods of the invention are useful for preventing retroviral infections				
CC	in cells susceptible to bacterial or viral infections, or treating				
CC	patients having infections such as HIV (human immunodeficiency virus),				
CC	sexually transmitted diseases, bacterial vaginosis or ophthalmic				
CC	infections. The retrocyclin-mediated killing is useful for modeling				
CC	and screening novel antibiotics. The invention is also useful in gene				
CC	therapy. The present sequence is rhesus monkey theta defensin 1b				
CC	precursor, RTD2 DNA. This sequence is used in the exemplification of				
CC	the invention.				
XX					
SQ	Sequence 495 BP; 103 A; 145 C; 122 G; 125 T; 0 other;				
Alignment Scores:					
Pred. No.:	0.00553	Length:	495		
Score:	83.00	Matches:	19		
Percent Similarity:	61.36%	Conservative:	8		
Best Local Similarity:	43.18%	Mismatches:	17		
Query Match:	35.62%	Indels:	0		
DB:	25	Gaps:	0		
US-10-045-180A-5 (1-44) x AAD51765 (1-495)					
QY	1 GlnuProLeuGlnAlaArgAlaHisGluMetProAlaGlnGlyGlnProProAlaAsp 20				
Db	150 GAGGCAACGTGAGGCAAGAGCTGATGAAGCTGCCGCCAGACAGACCTGAGCAGATGAT 209				
QY	21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40				
Db	210 CAGGGAATGGCTCATTCCTTTACACGCGCTGAAAGCGCGCTCTCCCTTCAGAGTCA 269				
QY	41 ThrIysGlyLeu 44				
Db	270 GCGAGAGGCTTG 281				
RESULT 13					
AAAC66924					
AAAC66924	standard; cDNA; 500 BP.				
AAC66924;					
27-MAR-2001 (first entry)					
Rhesus macaque theta defensin RTD-1a coding sequence.					
Theta defensin; antimicrobial; cyclic; bacterium; fungus; protozoan;					
virus; helminth; disinfectant; food preservative; analogue; ss.					
Rhesus macaque.					
WO200068265-A1.					
16-NOV-2000.					

```

XX 10-MAY-2000; 2000MO-US12842.
XX 10-MAY-1999; 99US-0309487.
XX (REGC ) UNIV CALIFORNIA.
XX Seisted ME, Tang Y, Yuan J, Ouellette AJ,
XX MPI; 2001-031853/04.
XX P-PSDB; AAB35040.
XX Novel theta defensin peptide with antimicrobial activity against
XX bacteria, yeast, fungi, protozoa and viruses -
XX Example 5; Fig 11; 110pp; English.
XX
XX The present invention provides theta defensin peptides and analogues
XX which have antimicrobial activity. They can be used in the treatment of
XX bacterial, viral, fungal, protozoan and helminthic infections, in
XX disinfectants and as food preservatives.
XX
SQ Sequence 500 BP; 110 A; 144 C; 115 G; 131 T; 0 other;

Alignment Scores:
Pred. No.: 0.00768 Length: 500
Score: 82.00 Matches: 19
Percent Similarity: 59.09% Conservative: 7
Best Local Similarity: 43.18% Mismatches: 18
Query Match: 35.19% Indels: 0
DB: 22 Gaps: 0

US-10-045-180A-5 (1-44) x AAC66924 (1-500)
QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAsp 20
Db 155 GAGGCGACGTGAGGCGAAGACTATATAGCTGCCGCCGACGACACTGGAGACGATGAT 214
QY 21 GlnAspValValIleTyrPheSerSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 215 CAGGGAATGCGCTCATCTCTTTACATGAGCGCTGAAAGCGCGCTTCCACTTTCAGATCA 274
QY 41 ThrIysGlyLeu 44
Db 275 GCGAAAGGCTTG 286

RESULT 14
AADS1768
ID AADS1768 standard; DNA; 500 BP.
XX
XX AADS1768;
XX
XX 16-APR-2003 (first entry)
XX
XX Macaca mulatta RTD3 DNA.
XX
XX Macaca mulatta RTD3 DNA.
XX
XX Retrocyclin; infection; sexually transmitted disease; gene therapy; HIV;
XX human immunodeficiency virus; bacterial vaginosis; ophthalmic infection;
XX antibiotic modelling; antimicrobial; rhesus monkey; theta defensin 1a;
XX RTD3; gene; ds.
XX
XX Macaca mulatta.
XX
XX Key Location/Qualifiers
XX FH 95..325
XX CDS
XX FT /*tag= a
XX FT /product= "RTD3 protein"
XX FT sig_peptide 95..154
XX FT /*tag= b
XX FT mat_peptide 155..322
XX FT /*tag= c
XX FT /product= "Mature RTD3 protein"
XX

```


PN W0200285401-A1.
XX
XX 31-OCT-2002.
PD
XX
XX 18-APR-2002; 2002WO-US12353.
PF
XX 18-APR-2001; 2001US-284855P.
PR
XX (REGC) UNIV CALIFORNIA.
PA
XX Lehrer RI, Waring AJ, Cole AM, Hong TB;
PI WPI; 2003-103387/09.
XX P-PSDB; AAE33812.
DR
XX
XX New isolated retrocyclin peptide, useful for preventing retroviral
PT infections in cells susceptible to bacterial or viral infections or
PT treating patients having the infections, such as HIV, sexually
PT transmitted diseases, vaginosis -
XX
XX Disclosure; Page 52; 72pp; English.
XX
XX The invention relates to novel retrocyclin peptides. Peptides and
CC methods of the invention are useful for preventing retroviral infections
CC in cells susceptible to bacterial or viral infections, or treating
CC patients having infections such as HIV (human immunodeficiency virus),
CC sexually transmitted diseases, bacterial vaginosis or ophthalmic
CC infections. The retrocyclin-mediated killing is useful for modelling
CC and screening novel antibiotics. The invention is also useful in gene
CC therapy. The present sequence is theus monkey theta defensin 1A
CC precursor, RND3 DNA. This sequence is used in the exemplification of
CC the invention.
XX
SQ Sequence 500 BP; 110 A; 144 C; 115 G; 131 T; 0 other;

Alignment Scores:
Pred. No.: 0.00788 Length: 500
Score: 82.00 Matches: 19
Percent Similarity: 59.09% Conservative: 7
Best Local Similarity: 43.18% Mismatches: 18
Query Match: 35.19% Indels: 0
DB: Gaps: 0

US-10-045-180A-5 (1-44) x AAD51768 (1-500)

QY 1 GIUProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
DB 155 GAGGCGACGTGACGAAAGAGCTGATGAGCTGCCGACGACGACCTTGGAACAGATGAT 214
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 215 CAGGGAATGGCTCATTCCTTTACATGAGCTGAAAGCGCGCTCTCCACTTCAAGATCA 274
QY 41 ThrIysGlyLeu 44
DB 275 GCGAAGAGCTTG 286

RESULT 15
AAQ53216
ID AAQ53216 standard; cDNA; 424 BP.
XX
XX AAQ53216;
AC
XX
DT 25-MAR-2003 (updated)
DT 17-JUN-1994 (first entry)
XX
XX Sequence encoding gastrointestinal defensin (GID) peptide called
DE human defensin 5.
XX
XX Gastrointestinal defensin peptide; GID; pharmaceutical; Paneth
KW cell; antimicrobial; anti-inflammatory; diagnosis;
KM contact disinfectant; ss.
XX

OS Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 10..294
FT CDS /*tag= a
FT
XX
XX M09324513-A1.
PN
XX
XX 09-DEC-1993.
PD
XX 18-MAY-1993; 93WO-US04740.
PF
XX 22-MAY-1992; 92US-0888232.
PR
XX (CHIL-) CHILDRENS HOSPITAL PHILADELPHIA.
PA
XX
XX Bevins CL, Jones DE;
PI
XX WPI; 1993-405719/50.
DR
XX P-PSDB; AAR44818.
DR
XX
XX Gastrointestinal defensin peptide(s) - useful as antimicrobial
PT and anti-inflammatory agents and for detecting gastrointestinal
PT disorders
XX
XX Claim 3; page 45-46; 97pp; English.
XX
XX A probe (D5' oligo) having the sequence given in AAQ53219 (claimed)
CC was used to screen a population of clones to identify novel defensin
CC defensin peptides. Pref. the libraries were human genomic and cDNA
CC libraries. Hybridisation and partial sequence analysis of the
CC identified clones contained previously characterised myeloid derived
CC defensin sequences as well as new defensin related sequences. Two
CC clones expressing new defensin related sequences were extensively
CC characterised and found to contain genes selectively expressed in
CC Paneth cells of the small intestine. These Paneth cell-derived
CC defensins are designated human defensin 5 and human defensin 6 and
CC are referred to as gastrointestinal defensin peptides. Comparison
CC of the deduced AA sequences of the defensin 5 cDNA with the
CC previously reported preprodefensin shows significant similarity.
CC The deduced AA sequence of defensin 6 cDNA has features similar to
CC defensin 5 and the previously reported preprodefensins.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 424 BP; 110 A; 118 C; 95 G; 101 T; 0 other;

Alignment Scores:
Pred. No.: 0.0127 Length: 424
Score: 80.00 Matches: 18
Percent Similarity: 60.00% Conservative: 6
Best Local Similarity: 45.00% Mismatches: 16
Query Match: 34.33% Indels: 0
DB: Gaps: 0

US-10-045-180A-5 (1-44) x AAQ53216 (1-424)

QY 1 GIUProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
DB 67 GAGTCACTCCAGAAAGAGCTGATGAGCTACACCCAGAAAGAGCTCGGGAACAC 126
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 127 CAGGACCTTGCTATCTCCTTGACAGAAATGAGACTCTGCTTAGAACCTCAAGTTCT 186

Search completed: December 17, 2003, 15:06:03
Job time : 166 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_pzn model

Run on: December 17, 2003, 15:06:18 ; Search time 152.128 Seconds
(without alignments)
963.776 Million cell updates/sec

Title: US-10-045-180A-5

Perfect score: 233
Sequence: 1 EPLARAHMPAKQKPPADD.....IYFGSDSCSLQVPGSTKGL 44

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2211978 seqs, 166101734 residues

Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=framer_pzn.model -DEV=xlp
-O=/cgn2_1/USPFO spool_p/US10045180/rnat_17122003_145026_12635/app_query.fasta_1.860
-DB=Published Applications_NA -QPMT=faacbp -SUFFIX=rnpb -WINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Diobum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORES=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10045180 @CGN 1.1 174 @rnat_17122003_145026_12635
-NCPU=6 -ICPU=3 -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications_NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/ECT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/ECTS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09B_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US09A_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARYS

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	233	100.0	453	14	US-10-013-770-2	Sequence 2, Appl1
2	233	100.0	453	14	US-10-045-180A-2	Sequence 2, Appl1
3	209	89.7	4415	14	US-10-013-770-1	Sequence 1, Appl1
4	209	89.7	4415	14	US-10-045-180A-1	Sequence 1, Appl1
5	114	48.9	542	14	US-10-045-180A-8	Sequence 8, Appl1
6	114	48.9	860	14	US-10-044-090-473	Sequence 473, Appl
7	104	44.6	496	13	US-10-141-645-11	Sequence 11, Appl
8	104	44.6	1348	15	US-10-125-237-78	Sequence 78, Appl
9	104	44.6	1348	15	US-10-105-891-78	Sequence 78, Appl
10	100	42.9	592	13	US-10-027-632-30053	Sequence 30053, A
11	100	42.9	592	13	US-10-027-632-30053	Sequence 30053, A
12	100	42.9	592	14	US-10-027-632-30053	Sequence 30053, A
13	100	42.9	592	14	US-10-027-632-30053	Sequence 30053, A
14	95	40.8	654	13	US-10-027-632-141529	Sequence 141529, A
15	95	40.8	654	14	US-10-027-632-141529	Sequence 141529, A
16	95	40.8	741	13	US-10-027-632-141528	Sequence 141528, A
17	95	40.8	741	13	US-10-027-632-141530	Sequence 141530, A
18	95	40.8	741	14	US-10-027-632-141528	Sequence 141528, A
19	95	40.8	741	14	US-10-027-632-141530	Sequence 141530, A
20	95	40.8	4295	14	US-10-045-180A-7	Sequence 7, Appl1
21	90	38.6	437	13	US-10-141-645-120	Sequence 120, App
22	90	38.6	437	13	US-10-141-645-122	Sequence 122, App
23	90	38.6	437	13	US-10-141-645-124	Sequence 124, App
24	86	36.9	243	13	US-10-313-994-28	Sequence 28, Appl
25	83	35.6	495	13	US-10-141-645-16	Sequence 16, Appl
26	83	35.6	495	13	US-10-313-994-15	Sequence 15, Appl
27	82	35.2	500	13	US-10-141-645-14	Sequence 14, Appl
28	82	35.2	500	13	US-10-313-994-13	Sequence 13, Appl
29	80	34.3	2553	13	US-10-313-994-25	Sequence 25, Appl
30	78.5	33.7	337	11	US-09-918-995-18158	Sequence 18158, A
31	78.5	33.7	337	11	US-09-918-995-19070	Sequence 19070, A
32	78.5	33.7	353	11	US-09-918-995-19070	Sequence 19070, A
33	78.5	33.7	555	13	US-10-252-157-223	Sequence 223, App
34	78.5	33.7	726	10	US-09-816-828-16	Sequence 16, Appl
35	76	32.6	2525	13	US-10-313-994-24	Sequence 24, Appl
36	69	29.6	412	10	US-09-736-457-1047	Sequence 1047, Ap
37	69	29.6	412	10	US-09-902-841-1047	Sequence 1047, Ap
38	69	29.6	412	10	US-09-849-626-1047	Sequence 1047, Ap
39	69	29.6	412	13	US-10-113-872-1047	Sequence 1047, Ap
40	69	29.6	412	15	US-10-017-754-1047	Sequence 1047, Ap
41	68.5	29.4	597	9	US-09-864-761-15342	Sequence 15342, A
42	68.5	29.4	598	13	US-10-027-632-48280	Sequence 48280, A
43	68.5	29.4	598	13	US-10-027-632-48281	Sequence 48281, A
44	68.5	29.4	598	13	US-10-027-632-48282	Sequence 48282, A
45	68.5	29.4	598	13	US-10-027-632-48283	Sequence 48283, A

ALIGNMENTS

RESULT 1
US-10-013-770-2
Sequence 2, Application US/10013770
Publication No. US20020115151A1

GENERAL INFORMATION:

APPLICANT: GENSET SA

TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC

COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND

THERAPEUTIC APPLICATIONS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESS: Knodbe, Martens, Olsson & Beaz

STREET: 550 West C Street

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92101

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/013,770

FILING DATE: 10-Dec-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/486,580
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.064C1
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 BASE PAIRS
TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-013-770-2

Alignment Scores:
Pred. No.: 9,49e-29 Length: 453
Score: 233.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-045-180a-5 (1-44) x US-10-013-770-2 (1-453)

QY 1 GIUProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAsp 20
DB 109 GAGCCGCTCCAGGCAAGAGCTCATGATGCCAGCCAGAGAGGCTCCAGAGATGAC 168
QY 21 GlnAspValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 169 CAGGATGTGCTATTACTTTTCAGAGATGACAGCTGCTCTTCCAGGTTCCAGGCTCA 228
QY 41 ThrIlySGlyLeu 44
DB 229 ACMAAGGCTTG 240

RESULT 2
US-10-045-180A-2
Sequence 2, Application US/10045180A
Publication No. US20020182703A1
GENERAL INFORMATION:
APPLICANT: Bouquelerec, Lydie
APPLICANT: Chumakov, Ilya
TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Compositi
FILE REFERENCE: GEN-100D1
CURRENT APPLICATION NUMBER: US/10/045,180A
CURRENT FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: US 09/486,580
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: PCT/FR98/01864
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: FR 97/10823
PRIOR FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 453
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (52)..(336)
OTHER INFORMATION: Def-X coding sequence
US-10-045-180A-2
Alignment Scores:

Pred. No.: 9,49e-29 Length: 453
Score: 233.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0
US-10-045-180A-5 (1-44) x US-10-045-180A-2 (1-453)

QY 1 GIUProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAsp 20
DB 109 GAGCCGCTCCAGGCAAGAGCTCATGATGCCAGCCAGAGAGGCTCCAGAGATGAC 168
QY 21 GlnAspValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 169 CAGGATGTGCTATTACTTTTCAGAGATGACAGCTGCTCTTCCAGGTTCCAGGCTCA 228
QY 41 ThrIlySGlyLeu 44
DB 229 ACMAAGGCTTG 240

RESULT 3
US-10-013-770-1
Sequence 1, Application US/10013770
Publication No. US20020115151A1
GENERAL INFORMATION:
APPLICANT: GENSET SA
TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
THERAPEUTIC APPLICATIONS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Knobbe, Martens, Olson & Bear
STREET: 550 West C Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/013,770
FILING DATE: 10-Dec-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/486,580
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.064C1
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4415 BASE PAIRS
TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: DNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Exon 1
LOCATION: 1836..1874
FEATURE:
NAME/KEY: Exon 2
LOCATION: 3394..3577
FEATURE:
NAME/KEY: Exon 3
LOCATION: 4161..4380
FEATURE:
NAME/KEY: start CDS

```

1 NAME/KEY: misc_feature
2 LOCATION: (1111)..(1111)
3 OTHER INFORMATION: n = a, c, g, or t.
4 FEATURE:
5 NAME/KEY: misc_feature
6 LOCATION: (1150)..(1150)
7 OTHER INFORMATION: n = a, c, g, or t.
8 FEATURE:
9 NAME/KEY: CAAT_signal
10 LOCATION: (1711)..(1714)
11 OTHER INFORMATION:
12 FEATURE:
13 NAME/KEY: TATA_signal
14 LOCATION: (1758)..(1767)
15 OTHER INFORMATION:
16 FEATURE:
17 NAME/KEY: misc_feature
18 LOCATION: (1780)..(1780)
19 OTHER INFORMATION: n = a, c, g, or t.
20 FEATURE:
21 NAME/KEY: misc_feature
22 LOCATION: (1836)..(1874)
23 OTHER INFORMATION: Exon 1
24 FEATURE:
25 NAME/KEY: misc_feature
26 LOCATION: (1875)..(1880)
27 OTHER INFORMATION: splice donor site
28 FEATURE:
29 NAME/KEY: misc_feature
30 LOCATION: (1974)..(1974)
31 OTHER INFORMATION: n = a, c, g, or t.
32 FEATURE:
33 NAME/KEY: misc_feature
34 LOCATION: (2117)..(2117)
35 OTHER INFORMATION: n = a, c, g, or t.
36 FEATURE:
37 NAME/KEY: misc_feature
38 LOCATION: (2133)..(2133)
39 OTHER INFORMATION: n = a, c, g, or t.
40 FEATURE:
41 NAME/KEY: misc_feature
42 LOCATION: (2155)..(2335)
43 OTHER INFORMATION: Alu insertion
44 FEATURE:
45 NAME/KEY: misc_feature
46 LOCATION: (2186)..(2186)
47 OTHER INFORMATION: n = a, c, g, or t.
48 FEATURE:
49 NAME/KEY: misc_feature
50 LOCATION: (2191)..(2191)
51 OTHER INFORMATION: n = a, c, g, or t.
52 FEATURE:
53 NAME/KEY: misc_feature
54 LOCATION: (2367)..(2367)
55 OTHER INFORMATION: n = a, c, g, or t.
56 FEATURE:
57 NAME/KEY: misc_feature
58 LOCATION: (2710)..(2780)
59 OTHER INFORMATION: L1 fragment insertion
60 FEATURE:
61 NAME/KEY: misc_feature
62 LOCATION: (3391)..(3393)
63 OTHER INFORMATION: splice acceptor site
64 FEATURE:
65 NAME/KEY: misc_feature
66 LOCATION: (3394)..(3577)
67 OTHER INFORMATION: Exon 2
68 FEATURE:
69 NAME/KEY: misc_feature
70 LOCATION: (3406)..(3408)
71 OTHER INFORMATION: Translation initiation codon (ATG)
72 FEATURE:
73 NAME/KEY: misc_feature

```

```
/ LOCATION: (3578).. (3583)
/ OTHER INFORMATION: splice donor site
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (4123).. (4123)
/ OTHER INFORMATION: n = a, c, g, or t.
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (4161).. (4163)
/ OTHER INFORMATION: splice acceptor site
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (4164).. (4379)
/ OTHER INFORMATION: Exon 3
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (4274).. (4276)
/ OTHER INFORMATION: Translation termination codon (TAA)
/ FEATURE:
/ NAME/KEY: polyA.signal
/ LOCATION: (4374).. (4379)
/ OTHER INFORMATION:
US-10-045-180A-1
```

```
Alignment Scores:
Pred. No.: 1.19e-23 Length: 4415
Score: 209.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.70% Indels: 0
DB: 14 Gaps: 0
```

US-10-045-180A-5 (1-44) x US-10-045-180A-1 (1-4415)

```
QY 1 GIUPProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlySGlnProProAlaAspAsp 20
DB 3463 GAGCCGCTCCAGGCAAGAGCTCATGAGATGCCAGCCAGACAGCTCCAGCAAGATGAC 3522
QY 21 GINAsPValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 39
DB 3523 CAGGATGTGCTATTACTTTTCAGAGATGACAGCTGCTCTTCAGGTTCCAGGT 3579
```

```
RESULT 5
US-10-045-180A-8
/ Sequence 8, Application US/10045180A
/ Publication No. US20020182703A1
/ GENERAL INFORMATION:
/ APPLICANT: Bougueleret, Lydie
/ TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Compositi
/ FILE REFERENCE: GEN-100D1
/ CURRENT APPLICATION NUMBER: US/10/045,180A
/ PRIOR FILING DATE: 2001-10-18
/ PRIOR APPLICATION NUMBER: US 09/486,580
/ PRIOR FILING DATE: 2000-02-25
/ PRIOR APPLICATION NUMBER: PCT/FR98/01864
/ PRIOR FILING DATE: 1998-08-28
/ PRIOR APPLICATION NUMBER: FR 97/10823
/ PRIOR FILING DATE: 1997-08-29
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 8
/ LENGTH: 542
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (52).. (345)
/ OTHER INFORMATION: Def-4 (HNP-4) coding sequence
US-10-045-180A-8
```

Alignment Scores:

```
Pred. No.: 3.26e-09 Length: 542
Score: 114.00 Matches: 23
Percent Similarity: 69.77% Conservative: 7
Best Local Similarity: 53.49% Mismatches: 13
Query Match: 48.93% Indels: 0
DB: 14 Gaps: 0
```

US-10-045-180A-5 (1-44) x US-10-045-180A-8 (1-542)

```
QY 2 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlySGlnProProAlaAspAspGln 21
DB 112 CCATCCAGGCAAGAGCTCATGAGATGCCAGCCAGACAGCTCCAGCAAGATGAC 171
QY 22 AsPValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 41
DB 172 GACATATCTATTCTCTTTCATGGGATTAAGCTCTGCTTCAGGTTTCAGGCTCAACA 231
QY 42 LysGlyLeu 44
DB 232 AGGGGCATG 240
```

```
RESULT 6
US-10-044-090-473
/ Sequence 473, Application US/10044090
/ Publication No. US20020137081A1
/ GENERAL INFORMATION:
/ APPLICANT: Olga Bandman
/ TITLE OF INVENTION: GENES DIFFERENTIALLY EXPRESSED IN VASCULAR TISSUE ACTIVATION
/ FILE REFERENCE: PA-0028 US
/ CURRENT APPLICATION NUMBER: US/10/044,090
/ PRIOR FILING DATE: 2002-01-09
/ NUMBER OF SEQ ID NOS: 850
/ SOFTWARE: PERL Program
/ SEQ ID NO 473
/ LENGTH: 860
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ OTHER INFORMATION: Incyte ID No. US20020137081A1 1670142CB1
US-10-044-090-473
```

```
Alignment Scores:
Pred. No.: 5.65e-09 Length: 860
Score: 114.00 Matches: 23
Percent Similarity: 69.77% Conservative: 7
Best Local Similarity: 53.49% Mismatches: 13
Query Match: 48.93% Indels: 0
DB: 14 Gaps: 0
```

US-10-045-180A-5 (1-44) x US-10-044-090-473 (1-860)

```
QY 2 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlySGlnProProAlaAspAspGln 21
DB 136 CCATCCAGGCAAGAGCTCATGAGATGCCAGCCAGACAGCTCCAGCAAGATGAC 195
QY 22 AsPValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySerThr 41
DB 196 GACATATCTATTCTCTTTCATGGGATTAAGCTCTGCTTCAGGTTTCAGGCTCAACA 255
QY 42 LysGlyLeu 44
DB 256 AGGGGCATG 264
```

```
RESULT 7
US-10-141-645-11
/ Sequence 11, Application US/10141645
/ Publication No. US20030144184A1
/ GENERAL INFORMATION:
/ APPLICANT: Robert Lehrer
/ APPLICANT: Alan Waring
/ APPLICANT: Alexander Cole
/ APPLICANT: Teresa Hong
```

```

; TITLE OF INVENTION: Retrocyclins - Antiviral and
; TITLE OF INVENTION: Antimicrobial Peptides
; FILE REFERENCE: UCLA-001CIP
; CURRENT APPLICATION NUMBER: US/10/141,645
; CURRENT FILING DATE: 2002-05-06
; PRIOR APPLICATION NUMBER: 60/284,855
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: Unassigned
; PRIOR FILING DATE: 2002-04-18
; NUMBER OF SEQ ID NOS: 125
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 496
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (124)...(304)
; OTHER INFORMATION: retrocyclin
US-10-141-645-11

Alignment Scores:
Pred. No.: 1,26e-07 Length: 496
Score: 104.00 Matches: 22
Percent Similarity: 68.18% Conservative: 8
Best Local Similarity: 50.00% Mismatches: 14
Query Match: 44.64% Indels: 0
DB: 13 Gaps: 0

US-10-045-180A-5 (1-44) x US-10-141-645-11 (1-496)

QY 1 GIUProLeuGlnAlaArgAlaHisGlnMetProAlaGlnIlySGlnProAlaAsp 20
    |||||
DB 136 GAGCCACTTCAGGCAAGAGCTGATGAGCTCGAGCCGAGCAGCTTGAGCAGATGAT 195
    |||||

QY 21 GlnAspValAlIleTyPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
    |||||
DB 196 CAGGAAATGGCTCATGCTTTACATGGCATGAAAGTCCGCTTCCTTCAGACTCA 255
    |||||

QY 41 ThrIlySGlyLeu 44
    |||||
DB 256 GCGAGAGGCTTG 267
    |||||

RESULT 8
US-10-125-237-78
; Sequence 78, Application US/10125237
; Publication No. US20030022329A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhang, Jie
; APPLICANT: Weinman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dimaac, Radoje T.
; TITLE OF INVENTION: No. US20030022329A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 791CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/125,237
; CURRENT FILING DATE: 2002-04-17
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Pt_Fl_genes Version 2.0
; SEQ ID NO 78
; LENGTH: 1348
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ..(1143)
US-10-125-237-78

Alignment Scores:
Pred. No.: 4.15e-07 Length: 1348
Score: 104.00 Matches: 23
Percent Similarity: 68.29% Conservative: 5
Best Local Similarity: 56.10% Mismatches: 13
Query Match: 44.64% Indels: 0
DB: 15 Gaps: 0

US-10-045-180A-5 (1-44) x US-10-125-237-78 (1-1348)

QY 1 GIUProLeuGlnAlaArgAlaHisGlnMetProAlaGlnIlySGlnProAlaAsp 20
    |||||
DB 421 GAGCCACTTCAGGCAAGAGCTGATGAGCTCGAGCCGAGCAGCTTGAGCAGATGAT 480
    |||||

QY 21 GlnAspValAlIleTyPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
    |||||
DB 481 CAGGAAATGGCTCATGCTTTGATGGCATGAAAGAGCTCCTTCAGCTTCAGGAGAG 540
    |||||

QY 41 Thr 41
    |||||
DB 541 TCC 543
    |||||

RESULT 9
US-10-105-891-78
; Sequence 78, Application US/10105891
; Publication No. US20030073099A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Zhou, Ping
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhang, Jie
; APPLICANT: Weinman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dimaac, Radoje T.
; TITLE OF INVENTION: No. US20030073099A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 791CIP2A
; CURRENT APPLICATION NUMBER: US/10/105,891
; CURRENT FILING DATE: 2002-03-25
; PRIOR APPLICATION NUMBER: 09/668,317
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 91
; SOFTWARE: Pt_Fl_genes Version 2.0
; SEQ ID NO 78
; LENGTH: 1348
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ..(1143)
US-10-105-891-78

Alignment Scores:
Pred. No.: 4.15e-07 Length: 1348
Score: 104.00 Matches: 23
Percent Similarity: 68.29% Conservative: 5
Best Local Similarity: 56.10% Mismatches: 13
Query Match: 44.64% Indels: 0
DB: 15 Gaps: 0

US-10-045-180A-5 (1-44) x US-10-105-891-78 (1-1348)
```

```
QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLySGlnProProAlaAspAsp 20
DB 421 GAGCCACTCCAGGCAAGTGTGATGAGCTTACAGCCAGAGAGAGCTGGAGCAGATGAT 480
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 481 CAGGAATGCTGCTCATGCTTTACATGAGCAATGAAGTCCGCTTCCGCTTTCAGGTGAG 540
QY 41 Thr 41
DB 541 TCC 543

RESULT 10
US-10-027-632-30053
; Sequence 30053, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30053
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(592)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-30053

Alignment Scores:
Pred. No.: 7,02e-07 Length: 592
Score: 100.00 Matches: 21
Percent Similarity: 68.29% Conservative: 7
Best Local Similarity: 51.22% Mismatches: 13
Query Match: 42.92% Indels: 0
DB: 13 Gaps: 0

US-10-045-180a-5 (1-44) x US-10-027-632-30053 (1-592)
QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLySGlnProProAlaAspAsp 20
DB 215 GAGCCACTCCAGGCAAGTGTGATGAGCTTACAGCCAGAGAGAGCTGGAGCAGATGAT 274
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 275 CAGGAATGCTGCTCATGCTTTACATGAGCAATGAAGTCCGCTTCCGCTTTCAGGTGAG 334
QY 41 Thr 41
DB 335 ACA 337

RESULT 11
```

```
US-10-027-632-30054
; Sequence 30054, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30054
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(592)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-30054

Alignment Scores:
Pred. No.: 7,02e-07 Length: 592
Score: 100.00 Matches: 21
Percent Similarity: 68.29% Conservative: 7
Best Local Similarity: 51.22% Mismatches: 13
Query Match: 42.92% Indels: 0
DB: 13 Gaps: 0

US-10-045-180a-5 (1-44) x US-10-027-632-30054 (1-592)
QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnLySGlnProProAlaAspAsp 20
DB 215 GAGCCACTCCAGGCAAGTGTGATGAGCTTACAGCCAGAGAGAGCTGGAGCAGATGAT 274
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 275 CAGGAATGCTGCTCATGCTTTACATGAGCAATGAAGTCCGCTTCCGCTTTCAGGTGAG 334
QY 41 Thr 41
DB 335 ACA 337

RESULT 12
US-10-027-632-30053
; Sequence 30053, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30053
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(592)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-30053
```

QY 2 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnIleArgGlnProProAlaAspArgGln 21

Db 96 CCACTCCAGGAGAGGTGATGAGGCTCCAGGCCAGAGACGCGTGAGGCCAAGACAG 155
Qy 22 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 156 GACATATCTATTCTTCCTTGCAATGAGATAAAGCTCTGCTTCAAGTTTCAGGT 209

RESULT 15

US-10-027-632-141529
; Sequence 141529, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 141529
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-141529

Alignment Scores:

Pred. No.:	5.18e-06	Length:	654
Score:	95.00	Matches:	20
Percent Similarity:	65.79%	Conservative:	5
Best Local Similarity:	52.63%	Mismatches:	13
Query Match:	40.77%	Indels:	0
DB:	14	Gaps:	0

US-10-045-180A-5 (1-44) x US-10-027-632-141529 (1-654)

Qy 2 ProLeuGlnAlaArgAlaHisGluMetProAlaGlnProAlaAspAspGln 21
Db 96 CCACTCCAGGAGAGGTGATGAGGCTCCAGGCCAGAGACGCGTGAGGCCAAGACAG 155
Qy 22 AspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 156 GACATATCTATTCTTCCTTGCAATGAGATAAAGCTCTGCTTCAAGTTTCAGGT 209

Search completed: December 17, 2003, 21:36:01
Job time : 154.128 secs

Gencore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 2308.83 seconds

(without alignments)
632.011 Million cell updates/sec

Title: US-10-045-180a-5

Sequence: 1 BFLQARAHMPAKQPPADD.....IVFSGDDSCSLVPGSTKGL 44

Scoring table:

BIOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 10.0 , Xgapext 0.5	
Xgapop 6.0 , Xgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 33363688 seqs, 16581889874 residues

Total number of hits satisfying chosen parameters: 66727376

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US10045180/runat.17122003.145025.12578/app.query.fasta.1.860
-DB=Pending Patents NA_Main -OFM=fastap -SUPFIX=trpm -MINMATCH=0.1 -HOOPT=0
-LOOPTXT=0 -UNITS=bits -STRIP=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptlo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180 @cgn 1.1 9409 @runat.17122003.145025.12578 -NCPU=6 -ICPU=3
-NO MMAP -IARGBOUTERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents NA_Main:

1: /cgn2_6/ptodata/2/pna/PCUTS_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/PCUTS_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US06_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US08_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US081_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US083_COMB.seq.*
9: /cgn2_6/ptodata/2/pna/US084_COMB.seq.*
10: /cgn2_6/ptodata/2/pna/US085_COMB.seq.*
11: /cgn2_6/ptodata/2/pna/US086_COMB.seq.*
12: /cgn2_6/ptodata/2/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/2/pna/US088_COMB.seq.*
14: /cgn2_6/ptodata/2/pna/US089_COMB.seq.*
15: /cgn2_6/ptodata/2/pna/US090_COMB.seq.*
16: /cgn2_6/ptodata/2/pna/US091_COMB.seq.*
17: /cgn2_6/ptodata/2/pna/US092A_COMB.seq.*
18: /cgn2_6/ptodata/2/pna/US092B_COMB.seq.*
19: /cgn2_6/ptodata/2/pna/US093A_COMB.seq.*
20: /cgn2_6/ptodata/2/pna/US093B_COMB.seq.*
21: /cgn2_6/ptodata/2/pna/US094_COMB.seq.*
22: /cgn2_6/ptodata/2/pna/US095A_COMB.seq.*
23: /cgn2_6/ptodata/2/pna/US095B_COMB.seq.*
24: /cgn2_6/ptodata/2/pna/US095C_COMB.seq.*
25: /cgn2_6/ptodata/2/pna/US095D_COMB.seq.*
26: /cgn2_6/ptodata/2/pna/US096A_COMB.seq.*
27: /cgn2_6/ptodata/2/pna/US096B_COMB.seq.*
28: /cgn2_6/ptodata/2/pna/US096C_COMB.seq.*

29: /cgn2_6/ptodata/2/pna/US096D_COMB.seq.*
30: /cgn2_6/ptodata/2/pna/US096E_COMB.seq.*
31: /cgn2_6/ptodata/2/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/2/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/2/pna/US097C_COMB.seq.*
34: /cgn2_6/ptodata/2/pna/US098A_COMB.seq.*
35: /cgn2_6/ptodata/2/pna/US098B_COMB.seq.*
36: /cgn2_6/ptodata/2/pna/US098C_COMB.seq.*
37: /cgn2_6/ptodata/2/pna/US098D_COMB.seq.*
38: /cgn2_6/ptodata/2/pna/US099A_COMB.seq.*
39: /cgn2_6/ptodata/2/pna/US099B_COMB.seq.*
40: /cgn2_6/ptodata/2/pna/US099C_COMB.seq.*
41: /cgn2_6/ptodata/2/pna/US099D_COMB.seq.*
42: /cgn2_6/ptodata/2/pna/US099E_COMB.seq.*
43: /cgn2_6/ptodata/2/pna/US099F_COMB.seq.*
44: /cgn2_6/ptodata/2/pna/US100A_COMB.seq.*
45: /cgn2_6/ptodata/2/pna/US100B_COMB.seq.*
46: /cgn2_6/ptodata/2/pna/US101A_COMB.seq.*
47: /cgn2_6/ptodata/2/pna/US101B_COMB.seq.*
48: /cgn2_6/ptodata/2/pna/US102A_COMB.seq.*
49: /cgn2_6/ptodata/2/pna/US102B_COMB.seq.*
50: /cgn2_6/ptodata/2/pna/US103A_COMB.seq.*
51: /cgn2_6/ptodata/2/pna/US103B_COMB.seq.*
52: /cgn2_6/ptodata/2/pna/US104A_COMB.seq.*
53: /cgn2_6/ptodata/2/pna/US104B_COMB.seq.*
54: /cgn2_6/ptodata/2/pna/US6000_COMB.seq.*
55: /cgn2_6/ptodata/2/pna/US6001_COMB.seq.*
56: /cgn2_6/ptodata/2/pna/US6002_COMB.seq.*
57: /cgn2_6/ptodata/2/pna/US6003_COMB.seq.*
58: /cgn2_6/ptodata/2/pna/US6004_COMB.seq.*
59: /cgn2_6/ptodata/2/pna/US6005_COMB.seq.*
60: /cgn2_6/ptodata/2/pna/US6006_COMB.seq.*
61: /cgn2_6/ptodata/2/pna/US6007_COMB.seq.*
62: /cgn2_6/ptodata/2/pna/US6008_COMB.seq.*
63: /cgn2_6/ptodata/2/pna/US6009_COMB.seq.*
64: /cgn2_6/ptodata/2/pna/US6010_COMB.seq.*
65: /cgn2_6/ptodata/2/pna/US6011_COMB.seq.*
66: /cgn2_6/ptodata/2/pna/US6012_COMB.seq.*
67: /cgn2_6/ptodata/2/pna/US6013_COMB.seq.*
68: /cgn2_6/ptodata/2/pna/US6014_COMB.seq.*
69: /cgn2_6/ptodata/2/pna/US6015_COMB.seq.*
70: /cgn2_6/ptodata/2/pna/US6016_COMB.seq.*
71: /cgn2_6/ptodata/2/pna/US6017_COMB.seq.*
72: /cgn2_6/ptodata/2/pna/US6018_COMB.seq.*
73: /cgn2_6/ptodata/2/pna/US6019_COMB.seq.*
74: /cgn2_6/ptodata/2/pna/US6020_COMB.seq.*
75: /cgn2_6/ptodata/2/pna/US6021_COMB.seq.*
76: /cgn2_6/ptodata/2/pna/US6022_COMB.seq.*
77: /cgn2_6/ptodata/2/pna/US6023_COMB.seq.*
78: /cgn2_6/ptodata/2/pna/US6024_COMB.seq.*
79: /cgn2_6/ptodata/2/pna/US6025_COMB.seq.*
80: /cgn2_6/ptodata/2/pna/US6026_COMB.seq.*
81: /cgn2_6/ptodata/2/pna/US6027_COMB.seq.*
82: /cgn2_6/ptodata/2/pna/US6028_COMB.seq.*
83: /cgn2_6/ptodata/2/pna/US6029_COMB.seq.*
84: /cgn2_6/ptodata/2/pna/US6030_COMB.seq.*
85: /cgn2_6/ptodata/2/pna/US6031_COMB.seq.*
86: /cgn2_6/ptodata/2/pna/US6032_COMB.seq.*
87: /cgn2_6/ptodata/2/pna/US6033_COMB.seq.*
88: /cgn2_6/ptodata/2/pna/US6034_COMB.seq.*
89: /cgn2_6/ptodata/2/pna/US6035_COMB.seq.*
90: /cgn2_6/ptodata/2/pna/US6036_COMB.seq.*
91: /cgn2_6/ptodata/2/pna/US6037_COMB.seq.*
92: /cgn2_6/ptodata/2/pna/US6038_COMB.seq.*
93: /cgn2_6/ptodata/2/pna/US6039_COMB.seq.*
94: /cgn2_6/ptodata/2/pna/US6040_COMB.seq.*
95: /cgn2_6/ptodata/2/pna/US6041_COMB.seq.*
96: /cgn2_6/ptodata/2/pna/US6042_COMB.seq.*
97: /cgn2_6/ptodata/2/pna/US6043_COMB.seq.*
98: /cgn2_6/ptodata/2/pna/US6044_COMB.seq.*
99: /cgn2_6/ptodata/2/pna/US6045_COMB.seq.*
100: /cgn2_6/ptodata/2/pna/US6046_COMB.seq.*
101: /cgn2_6/ptodata/2/pna/US6047_COMB.seq.*

102: /cgn2_6/prodata/2/pna/US6047_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	233	100.0	285	47	US-10-170-235-6564	Sequence 6564, Ap
2	233	100.0	453	44	US-10-013-770-2	Sequence 2, Appl
3	233	100.0	453	44	US-10-045-180a-2	Sequence 2, Appl
4	209	89.7	4415	44	US-10-013-770-1	Sequence 1, Appl
5	209	89.7	4415	44	US-10-045-180a-1	Sequence 1, Appl
6	200	85.8	601	40	US-09-947-907-9094	Sequence 9094, Ap
7	200	85.8	610	70	US-60-160-203-825	Sequence 825, Ap
8	200	85.8	610	70	US-60-169-840-3181	Sequence 3181, Ap
9	200	85.8	696	75	US-60-160-203-156	Sequence 156, App
10	200	85.8	111999	75	US-60-212-664-312	Sequence 312, App
11	187	80.3	345	71	US-60-171-481-272	Sequence 272, App
12	187	80.3	345	71	US-60-171-481-272	Sequence 272, App
13	120	51.5	338	19	US-09-321-214-10452	Sequence 10452, A
14	120	51.5	338	22	US-09-516-335-10452	Sequence 10452, A
15	120	51.5	338	22	US-09-733-811-10452	Sequence 10452, A
16	120	51.5	338	32	US-09-733-811A-10452	Sequence 10452, A
17	120	51.5	338	42	US-09-975-640-10452	Sequence 10452, A
18	120	51.5	338	42	US-09-975-640A-10452	Sequence 10452, A
19	118	50.6	338	21	US-09-471-275-1730	Sequence 1730, Ap
20	118	50.6	338	21	US-09-488-725B-6871	Sequence 6871, Ap
21	114	48.9	271	16	US-09-179-473-1553	Sequence 1553, Ap
22	114	48.9	271	19	US-09-328-351-1553	Sequence 1553, Ap
23	114	48.9	271	30	US-09-359-067-26554	Sequence 26554, A
24	114	48.9	271	35	US-09-845-487-1553	Sequence 1553, Ap
25	114	48.9	271	35	US-09-845-487A-1553	Sequence 1553, Ap
26	114	48.9	304	18	US-09-271-490-1874	Sequence 1874, Ap
27	114	48.9	304	39	US-09-925-552-1874	Sequence 1874, Ap
28	114	48.9	304	44	US-10-032-354-1874	Sequence 1874, Ap
29	114	48.9	304	19	US-09-321-214-1874	Sequence 1874, Ap
30	114	48.9	349	14	US-09-516-335-3560	Sequence 3560, Ap
31	114	48.9	349	32	US-09-733-811-3560	Sequence 3560, Ap
32	114	48.9	349	32	US-09-733-811A-3560	Sequence 3560, Ap
33	114	48.9	349	42	US-09-975-640-3560	Sequence 3560, Ap
34	114	48.9	349	42	US-09-975-640A-3560	Sequence 3560, Ap
35	114	48.9	350	18	US-09-271-490-2574	Sequence 2574, Ap
36	114	48.9	350	39	US-09-925-552-2574	Sequence 2574, Ap
37	114	48.9	350	44	US-10-032-354-2574	Sequence 2574, Ap
38	114	48.9	382	17	US-09-234-611-10800	Sequence 10800, A
39	114	48.9	382	18	US-09-289-768-15603	Sequence 15603, A
40	114	48.9	382	38	US-09-904-809-10800	Sequence 10800, A
41	114	48.9	382	39	US-09-934-397-15603	Sequence 15603, A
42	114	48.9	396	18	US-09-289-768-13222	Sequence 13222, A
43	114	48.9	396	39	US-09-939-397-13222	Sequence 13222, A
44	114	48.9	521	32	US-09-726-811-171	Sequence 171, App
45	114	48.9	524	34	US-09-823-241-2013	Sequence 2013, Ap

ALIGNMENTS

RESULT 1
US-10-170-235-6564
Sequence 6564, Application US/10170235
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig
TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMAN TRANSCRIPTS, FOR DETECTING EXPRESSION AND OTHER USES THEREOF
FILE REFERENCE: CL001380
CURRENT APPLICATION NUMBER: US/10/170, 235
CURRENT FILING DATE: 2003-03-17
NUMBER OF SEQ ID NOS: 42514
SEQ ID NO 6564
LENGTH: 285

TYPE: DNA
ORGANISM: HUMAN
US-10-170-235-6564

Alignment Scores:

Pred. No.: 1.25e-22 Length: 285
Score: 233.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 47 Gaps: 0

US-10-045-180a-5 (1-44) x US-10-170-235-6564 (1-285)

QY 1 GUAProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProAlaAasp 20
DB 58 GACCCCTCCAGGCAAGACTATGATGCCGCCAAGAGAGCTTCAGAGATGAC 117
QY 21 GlnAspValIleTyrPheSerGlyAaspAspSerCysSerLeuGlnValProGlySer 40
DB 118 CAGAGATGTGCTATTACTTTTCAGAGATGACAGCTGCTCTTCAAGTTCAGGCTCA 177
QY 41 ThrIlyGlyLeu 44
DB 178 ACAAGAGGCTTG 189

RESULT 2
US-10-013-770-2
Sequence 2, Application US/10013770
GENERAL INFORMATION:

APPLICANT: GENSET SA

TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
THERAPEUTIC APPLICATIONS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS: Martens, Olson & Bear

ADDRESS: 550 West C Street
CITY: San Diego
STATE: California
COUNTRY: USA

ZIP: 92101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatibleOPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/013,770
FILING DATE: 10-Dec-2001PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/486,580
FILING DATE: <Unknown>ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637REFERENCE/DOCKET NUMBER: GENSET.064C1
INFORMATION FOR SEQ ID NO: 2:SEQUENCE CHARACTERISTICS:
LENGTH: 453 BASE PAIRSTYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLETOPOLOGY: LINEAR
MOLECULE TYPE: cDNAORIGINAL SOURCE:
ORGANISM: Homo sapiensSEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-013-770-2

Alignment Scores:

Pred. No.: 2.15e-22 Length: 453
Score: 233.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 44 Gaps: 0
US-10-045-180A-5 (1-44) x US-10-013-770-2 (1-453)

QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnIleGlnProProAlaAspAsp 20
Db 109 GAGCCGCTCCAGGCAAGAGCTCATGATGCCAGCCCAAGAGAGGCTCCAGCAGATGAC 168

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 169 CAGGATGTGTCATTACTTTTCAGGAGATGACAGCTGCTCTTCAGGTTCCAGGCTCA 228

QY 41 ThrIysGlyLeu 44
Db 229 ACAAGAGGCTTG 240

RESULT 3
US-10-045-180A-2
; Sequence 2, Application US/10045180A
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Compositi
; FILE REFERENCE: GEN-100D1
; CURRENT APPLICATION NUMBER: US/10/045,180A
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: US 09/486,580
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: PCT/FR98/01864
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: FR 97/10823
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 453
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52) ..(336)
; OTHER INFORMATION: Def-X coding sequence
US-10-045-180A-2

Alignment Scores:
Pred. No.: 2,15e-22 Length: 453
Score: 233.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 44 Gaps: 0
US-10-045-180A-5 (1-44) x US-10-045-180A-2 (1-453)

QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnIleGlnProProAlaAspAsp 20
Db 109 GAGCCGCTCCAGGCAAGAGCTCATGATGCCAGCCCAAGAGAGGCTCCAGCAGATGAC 168

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 169 CAGGATGTGTCATTACTTTTCAGGAGATGACAGCTGCTCTTCAGGTTCCAGGCTCA 228

QY 41 ThrIysGlyLeu 44
Db 229 ACAAGAGGCTTG 240

RESULT 4
US-10-013-770-1
; Sequence 1, Application US/10013770
; GENERAL INFORMATION:

APPLICANT: GENSET SA
TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
THERAPEUTIC APPLICATIONS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 550 West C Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/013,770
FILING DATE: 10-Dec-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/486,580
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.064C1
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4415 BASE PAIRS
TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: DNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Exon 1
LOCATION: 1836..1874
FEATURE:
NAME/KEY: Exon 2
LOCATION: 3394..3577
FEATURE:
NAME/KEY: Exon 3
LOCATION: 4161..4380
FEATURE:
NAME/KEY: start CDS
LOCATION: 3406..3408
FEATURE:
NAME/KEY: stop CDS
LOCATION: 4276..4278
FEATURE:
NAME/KEY: polyadenylation site
LOCATION: 4374..4379
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-013-770-1

Alignment Scores:
Pred. No.: 7e-18 Length: 4415
Score: 209.00 Matches: 39
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.70% Indels: 0
DB: 44 Gaps: 0
US-10-045-180A-5 (1-44) x US-10-013-770-1 (1-4415)

QY 1 GluProLeuGlnAlaArgAlaHisGluMetProAlaGlnIleGlnProProAlaAspAsp 20
Db 3463 GAGCCGCTCCAGGCAAGAGCTCATGATGCCAGCCCAAGAGAGGCTCCAGCAGATGAC 3522

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39

Db 3523 CAGGATGTGCTACTTTCTTTCAGAGATGACAGCTCTCTCTTTCAGTTCCAGGT 3579

RESULT 5
US-10-045-180a-1
Sequence 1, Application US/10045180A
GENERAL INFORMATION:
APPLICANT: Bougueleret, Lydie
APPLICANT: Chumakov, Ilya
TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Compositions
FILE REFERENCE: GEN-100D1
CURRENT APPLICATION NUMBER: US/10/045,180A
CURRENT FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: US 09/486,580
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: PCT/FR98/01864
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: FR 97/10823
PRIOR FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 4415
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)..(4415)
OTHER INFORMATION: Def-X genomic sequence
FEATURE:
NAME/KEY: misc.feature
LOCATION: (85)..(85)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (143)..(143)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (670)..(670)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (970)..(970)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1111)..(1111)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1150)..(1150)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: CAAT_signal
LOCATION: (1711)..(1714)
OTHER INFORMATION:
FEATURE:
NAME/KEY: TATA_signal
LOCATION: (1758)..(1767)
OTHER INFORMATION:
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1780)..(1780)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1836)..(1874)
OTHER INFORMATION: Exon 1
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1875)..(1880)
OTHER INFORMATION: splice donor site

FEATURE:
NAME/KEY: misc.feature
LOCATION: (1974)..(1974)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2117)..(2117)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2133)..(2133)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2155)..(2335)
OTHER INFORMATION: Alu insertion
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2186)..(2186)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2191)..(2191)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2367)..(2367)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (2710)..(2780)
OTHER INFORMATION: L1 fragment insertion
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3391)..(3393)
OTHER INFORMATION: splice acceptor site
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3395)..(3577)
OTHER INFORMATION: Exon 2
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3406)..(3408)
OTHER INFORMATION: Translation initiation codon (ATG)
FEATURE:
NAME/KEY: misc.feature
LOCATION: (3578)..(3583)
OTHER INFORMATION: splice donor site
FEATURE:
NAME/KEY: misc.feature
LOCATION: (4123)..(4123)
OTHER INFORMATION: n = a, c, g, or t.
FEATURE:
NAME/KEY: misc.feature
LOCATION: (4161)..(4163)
OTHER INFORMATION: splice acceptor site
FEATURE:
NAME/KEY: misc.feature
LOCATION: (4164)..(4379)
OTHER INFORMATION: Exon 3
FEATURE:
NAME/KEY: misc.feature
LOCATION: (4274)..(4276)
OTHER INFORMATION: Translation termination codon (TAA)
FEATURE:
NAME/KEY: polyA_signal
LOCATION: (4374)..(4379)
OTHER INFORMATION:
US-10-045-180a-1

Alignment Scores:
Pred. No.: 7e-18 Length: 4415
Score: 209.00 Matches: 39

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 89.70% Indels: 0
DB: 44 Gaps: 0

US-10-045-180A-5 (1-44) x US-10-045-180A-1 (1-4415)

QY 1 GiupProLeuglnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
DB 3463 GAGCCGCTCCAGGCAAGAGCTGATGATGCCAGCCCAAGACAGCTCCAGACATGAC 3522

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
DB 3523 CAGGATGGTGCATTACTTTTCAGAGATGACAGCTGCTCTCTTCAGATTCCAGGT 3579

RESULT 6
US-09-947-907-9094/c
; Sequence 9094, Application US/09947907
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: COLLECTION OF CODING REGION SINGLE
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISM (GSPs) LOCATED ON EACH OF THE HUMAN
; FILE REFERENCE: CHROMOSOMES, METHODS OF DETECTION, AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/947,907
; PRIOR FILING DATE: 2003-03-17
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 2126
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9094
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-947-907-9094

Alignment Scores:
Pred. No.: 1,24e-17 Length: 601
Score: 200.00 Matches: 38
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 1
Query Match: 85.84% Indels: 0
DB: 40 Gaps: 0

US-10-045-180A-5 (1-44) x US-09-947-907-9094 (1-601)

QY 1 GiupProLeuglnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
DB 322 GAGCCGCTCCAGGCAAGAGCTGATGATGCCAGCCCAAGACAGCTCCAGACATGAC 263

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
DB 263 CAGGATGGTGCATTACTTTTCAGAGATGACAGCTGCTCTCTTCAGATTCCAGGT 206

RESULT 7
US-60-160-203-825/c
; Sequence 825, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CLO00116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 825
; LENGTH: 610

TYPR: DNA
; ORGANISM: HUMAN
US-60-160-203-825

Alignment Scores:
Pred. No.: 1,26e-17 Length: 610
Score: 200.00 Matches: 38
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 1
Query Match: 85.84% Indels: 0
DB: 70 Gaps: 0

US-10-045-180A-5 (1-44) x US-60-160-203-825 (1-610)

QY 1 GiupProLeuglnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
DB 356 GAGCCGCTCCAGGCAAGAGCTGATGATGCCAGCCCAAGACAGCTCCAGACATGAC 297

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
DB 296 CAGGATGGTGCATTACTTTTCAGAGATGACAGCTGCTCTCTTCAGATTCCAGGT 240

RESULT 8
US-60-169-840-3181/c
; Sequence 3181, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CLO00164
; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 9628
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3181
; LENGTH: 610
; TYPE: DNA
; ORGANISM: Human
US-60-169-840-3181

Alignment Scores:
Pred. No.: 1,26e-17 Length: 610
Score: 200.00 Matches: 38
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 1
Query Match: 85.84% Indels: 0
DB: 70 Gaps: 0

US-10-045-180A-5 (1-44) x US-60-169-840-3181 (1-610)

QY 1 GiupProLeuglnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
DB 356 GAGCCGCTCCAGGCAAGAGCTGATGATGCCAGCCCAAGACAGCTCCAGACATGAC 297

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
DB 296 CAGGATGGTGCATTACTTTTCAGAGATGACAGCTGCTCTCTTCAGATTCCAGGT 240

RESULT 9
US-60-160-203-156
; Sequence 156, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CLO00116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 156

```

; LENGTH: 696
; TYPE: DNA
; ORGANISM: HUMAN
US-60-160-203-156

Alignment Scores:
Pred. No.: 1.47e-17 Length: 696
Score: 200.00 Matches: 38
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 1
Query Match: 85.84% Indels: 0
DB: 70 Gaps: 0

US-10-045-180A-5 (1-44) x US-60-160-203-156 (1-696)

QY 1 GIUProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAsp 20
Db 415 GAGCGGCTCCAGGCAAGAGCTGATGAGATGCCAGCCCAAGAGAGCTCCAGCAGATGAC 474

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 475 CAGGATGTGTCTATTACTTTTCAGAGATGACAGCTGCTCTTCCAGGTTCCAGGT 531

RESULT 10
US-60-212-664-312
; Sequence 312, Application US/60212664
; GENERAL INFORMATION:
; APPLICANT: Ladunga, Steve
; APPLICANT: Spier, Gene
; APPLICANT: Greenberg, Simon
; APPLICANT: Rabkin, Steven
; APPLICANT: Wang, Yu
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000687
; CURRENT APPLICATION NUMBER: US/60/212,664
; CURRENT FILING DATE: 2000-06-20
; NUMBER OF SEQ ID NOS: 636
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312
; LENGTH: 111999
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(111999)
; OTHER INFORMATION: n = A,T,C or G
US-60-212-664-312

Alignment Scores:
Pred. No.: 5.55e-15 Length: 111999
Score: 200.00 Matches: 38
Percent Similarity: 97.44% Conservative: 0
Best Local Similarity: 97.44% Mismatches: 1
Query Match: 85.84% Indels: 0
DB: 75 Gaps: 0

US-10-045-180A-5 (1-44) x US-60-212-664-312 (1-111999)

QY 1 GIUProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAsp 20
Db 67705 GAGCGGCTCCAGGCAAGAGCTGATGAGATGCCAGCCCAAGAGAGCTCCAGCAGATGAC 67764

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 67765 CAGGATGTGTCTATTACTTTTCAGAGATGACAGCTGCTCTTCCAGGTTCCAGGT 67821

RESULT 11
US-60-171-481-272/c
; Sequence 272, Application US/60171481
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
```

```

; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: US$ THEREOF
; FILE REFERENCE: CL000169
; CURRENT APPLICATION NUMBER: US/60/171,481
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 1898
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 272
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Human
US-60-171-481-272

Alignment Scores:
Pred. No.: 4.25e-16 Length: 345
Score: 187.00 Matches: 38
Percent Similarity: 95.00% Conservative: 0
Best Local Similarity: 95.00% Mismatches: 1
Query Match: 80.26% Indels: 1
DB: 71 Gaps: 0

US-10-045-180A-5 (1-44) x US-60-171-481-272 (1-345)

QY 1 GIUProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAsp 20
Db 219 GAGCGGCTCCAGGCAAGAGCTGATGAGATGCCAGCCCAATAGAGCTCCAGCAGATGA 160

QY 20 pGlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 159 CCAGATGTGTCTATTACTTTTCAGAGATGACAGCTGCTCTTCCAGGTTCCAGGT 102

RESULT 12
US-60-171-481-777/c
; Sequence 777, Application US/60171481
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000169
; CURRENT APPLICATION NUMBER: US/60/171,481
; CURRENT FILING DATE: 1999-12-22
; NUMBER OF SEQ ID NOS: 1898
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 777
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Human
US-60-171-481-777

Alignment Scores:
Pred. No.: 4.25e-16 Length: 345
Score: 187.00 Matches: 38
Percent Similarity: 95.00% Conservative: 0
Best Local Similarity: 95.00% Mismatches: 1
Query Match: 80.26% Indels: 1
DB: 71 Gaps: 0

US-10-045-180A-5 (1-44) x US-60-171-481-777 (1-345)

QY 1 GIUProLeuGlnAlaArgAlaHisGluMetProAlaGlnLysGlnProProAlaAspAsp 20
Db 219 GAGCGGCTCCAGGCAAGAGCTGATGAGATGCCAGCCCAATAGAGCTCCAGCAGATGA 160

QY 20 pGlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGly 39
Db 159 CCAGATGTGTCTATTACTTTTCAGAGATGACAGCTGCTCTTCCAGGTTCCAGGT 102

RESULT 13
US-09-321-214-10452/c
; Sequence 10452, Application US/09321214
; GENERAL INFORMATION:
```

```

; APPLICANT: Hyseq, Inc.
; APPLICANT: Drmanac, Radoje
; APPLICANT: Cirkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Ford, John
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 20411-740
; CURRENT APPLICATION NUMBER: US/09/321,214
; CURRENT FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: US 60/088,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10452
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-321-214-10452
```

```

Alignment Scores:
Pred. No.: 9.74e-07 Length: 338
Score: 120.00 Matches: 26
Percent Similarity: 72.73% Conservative: 6
Best Local Similarity: 59.09% Mismatches: 12
Query Match: 51.50% Indels: 0
DB: 19 Gaps: 0
```

US-10-045-180a-5 (1-44) x US-09-321-214-10452 (1-338)

```

QY 1 GIUPProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
    |||||
DB 260 GAGCCATCCAGGCAATTCCTGATGAGTTACAGCCGAGCAGCTTGAGCAATGAT 201
    |||||
QY 21 GlnAspValValIleTyRPheserGlyAspAspSerCysSerLeuGlnValProGlySer 40
    |||||
DB 200 CAGGAAGTGTGATTCCTTGATGATGATGAAGAAGCTCCTTCAGGTTTCAGGCTCA 141
    |||||
QY 41 ThrIlyGlyLeu 44
    |||||
DB 140 AGGAGGGGCTTG 129
    |||||
```

RESULT 14

US-09-516-335-10452/c
; Sequence 10452, Application US/09516335
; GENERAL INFORMATION:

```

; APPLICANT: Arteburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lyne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
```

```

; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Farida
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjwan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verma, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/516,335
; CURRENT FILING DATE: 2000-03-01
; EARLIER APPLICATION NUMBER: 09/321,214
; EARLIER FILING DATE: 1999-05-26
; EARLIER APPLICATION NUMBER: 60/088,041
; EARLIER FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10452
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-516-335-10452
```

```

Alignment Scores:
Pred. No.: 9.74e-07 Length: 338
Score: 120.00 Matches: 26
Percent Similarity: 72.73% Conservative: 6
Best Local Similarity: 59.09% Mismatches: 12
Query Match: 51.50% Indels: 0
DB: 22 Gaps: 0
```

US-10-045-180a-5 (1-44) x US-09-516-335-10452 (1-338)

```

QY 1 GIUPProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
    |||||
DB 260 GAGCCATCCAGGCAATTCCTGATGAGTTACAGCCGAGCAGCTTGAGCAATGAT 201
    |||||
QY 21 GlnAspValValIleTyRPheserGlyAspAspSerCysSerLeuGlnValProGlySer 40
    |||||
DB 200 CAGGAAGTGTGATTCCTTGATGATGATGAAGAAGCTCCTTCAGGTTTCAGGCTCA 141
    |||||
QY 41 ThrIlyGlyLeu 44
    |||||
DB 140 AGGAGGGGCTTG 129
    |||||
```

RESULT 15

US-09-733-811-10452/c
; Sequence 10452, Application US/09733811
; GENERAL INFORMATION:

```

; APPLICANT: Arteburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
```



```

; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjivan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR FILING DATE: 09/321,214
; PRIOR FILING DATE: 1999-05-26
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10452
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-733-811-10452

```

```

Alignment Scores:
Pred. No.:          9.74e-07      Length:      338
Score:             120.00         Matches:     26
Percent Similarity: 72.73%        Conservative: 6
Best Local Similarity: 59.09%      Mismatches:  12
Query Match:       51.50%         Indels:      0
DB:                32            Gaps:         0

```

US-10-045-180a-5 (1-44) x US-09-733-811-10452 (1-338)

```

QY      1  GlnuProleuGlnAlaArgAlaHisGluMetProAlaGlnGlnProAlaAspASP 20
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      260 GAGCCATCTCCAGCAATTGCTGATGAGGTTACAGCCAGAGCACTGAGAGCAGATGAT 201
QY      21  GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
         ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      200 CAGGAAGTGGTGTGATTCCTTTCATGAGGATGAAGAAGCTCCTCTTCAGGTTTCAGGCTCA 141
QY      41  ThrIysGlyLeu 44
         ::||| |||
Db      140 AGGAGGGGCTTG 129

```

Search completed: December 17, 2003, 21:11:01
 Job time : 2320.08 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 44 Seconds
(without alignments)
441.383 Million cell updates/sec

Title: US-10-045-180A-5

Perfect score: 233
Sequence: 1 EPLQARAHMPAQKQPPADD.....IYFGSDSDCSIQVPSRYGL 44

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n_model -DEV=xlp
-O=/cgn2_1/USPTO.spool.p/US10045180/runat.17122003.145024.12553/app_query.fasta_1.860
-DB=Issued Patents NA -QRTT=fastap -SUFFIX=rmi -MINMATCH=0.1 -LOOPL=0
-IOOPEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX=BioSum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=plco -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10045180 @CGN 1.1 193 @runat.17122003.145024.12553 -MCPU=6 -ICPU=3
-NO MAP -LARGEROVER_NEG SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCUS.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	233	100.0	453	4	US-09-486-580A-2
2	209	89.7	4415	4	US-09-486-580A-1
3	101	43.3	452	1	US-08-158-189-6
4	101	43.3	452	4	US-09-016-434-1198
5	86	36.9	243	4	US-09-309-487-28
6	86	36.9	243	4	US-09-967-808-28
7	83	35.6	495	4	US-09-309-487-15
8	83	35.6	495	4	US-09-967-808-15
9	82	35.2	500	4	US-09-309-487-13
10	82	35.2	500	4	US-09-967-808-13
11	80	34.3	424	1	US-08-158-189-4
12	80	34.3	2553	4	US-09-309-487-25

13	80	34.3	2553	4	US-09-967-808-25	Sequence 25, Appl
14	76	32.6	2525	4	US-09-309-487-24	Sequence 24, Appl
15	76	32.6	2525	4	US-09-967-808-24	Sequence 24, Appl
16	76	32.6	2880	1	US-08-158-189-1	Sequence 1, Appl
17	74	31.8	445	1	US-08-486-013-66	Sequence 66, Appl
18	74	31.8	445	2	US-08-482-279-66	Sequence 66, Appl
19	74	31.8	445	2	US-08-342-268-66	Sequence 66, Appl
20	74	31.8	445	3	US-09-015-968-66	Sequence 66, Appl
21	74	31.8	445	4	US-09-397-386-66	Sequence 66, Appl
22	74	31.8	2408	1	US-08-486-013-69	Sequence 69, Appl
23	74	31.8	2408	2	US-08-482-279-69	Sequence 69, Appl
24	74	31.8	2408	2	US-08-342-268-69	Sequence 69, Appl
25	74	31.8	2408	3	US-09-015-968-69	Sequence 69, Appl
26	74	31.8	2408	4	US-09-397-386-69	Sequence 69, Appl
27	72.5	31.1	218	1	US-08-158-189-38	Sequence 38, Appl
28	72	30.9	2457	1	US-08-486-013-68	Sequence 68, Appl
29	72	30.9	2457	2	US-08-482-279-68	Sequence 68, Appl
30	72	30.9	2457	2	US-08-342-268-68	Sequence 68, Appl
31	72	30.9	2457	3	US-09-015-968-68	Sequence 68, Appl
32	72	30.9	2457	4	US-09-397-386-68	Sequence 68, Appl
33	72	30.9	2551	1	US-08-486-013-70	Sequence 70, Appl
34	72	30.9	2551	2	US-08-482-279-70	Sequence 70, Appl
35	72	30.9	2551	2	US-08-342-268-70	Sequence 70, Appl
36	72	30.9	2551	3	US-09-015-968-70	Sequence 70, Appl
37	72	30.9	2551	4	US-09-397-386-70	Sequence 70, Appl
38	71	30.5	445	1	US-08-486-013-65	Sequence 65, Appl
39	71	30.5	445	2	US-08-482-279-65	Sequence 65, Appl
40	71	30.5	445	2	US-08-342-268-65	Sequence 65, Appl
41	71	30.5	445	2	US-08-482-279-67	Sequence 67, Appl
42	71	30.5	445	2	US-08-342-268-67	Sequence 67, Appl
43	71	30.5	445	2	US-08-482-279-67	Sequence 67, Appl
44	71	30.5	445	3	US-09-015-968-65	Sequence 65, Appl
45	71	30.5	445	3	US-09-397-386-67	Sequence 67, Appl

ALIGNMENTS

RESULT 1
US-09-486-580A-2
Sequence 2, Application US/09486580A
Patent No. 6329340
GENERAL INFORMATION:
APPLICANT: GENSET SA
TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
TITLE OF INVENTION: THERAPEUTIC APPLICATIONS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Knobbe, Martens, Olson & Bear
STREET: 550 West C Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/486,580A
FILING DATE: FEBRUARY 25, 2000
ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.064C1
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 BASE PAIRS
TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA

```

; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-486-580A-2

Alignment Scores:
Pred. No.: 6,24e-26      Length: 453
Score: 233.00           Matches: 44
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 100.00%             Indels: 0
DB: 4                     Gaps: 0

US-10-045-180A-5 (1-44) x US-09-486-580A-2 (1-453)

QY 1 GlnProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAspAsp 20
Db 109 GAGCCGCTCCAGGCAAGAGCTCATGAGATGCCAGCCAGAGAGCCTCCAGAGATGAC 168

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
Db 169 CAGATGTGCTCATTTACTTTTCAGAGATGACAGCTGCTCTTCAGGTTCCAGGCTCA 228

QY 41 ThrIlySGlyLeu 44
Db 229 ACAAGGCGCTTG 240

RESULT 2
US-09-486-580A-1
; Sequence 1, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30 (EBO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4415 BASE PAIRS
; TYPE: NUCLEOTIDE
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: DNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Exon 1
; LOCATION: 1836..1874
; FEATURE:
; NAME/KEY: Exon 2
; LOCATION: 3394..3577
; FEATURE:
; NAME/KEY: Exon 3
; LOCATION: 4161..4380
```

```

; FEATURE:
; NAME/KEY: start CDS
; LOCATION: 3406..3408
; FEATURE:
; NAME/KEY: stop CDS
; LOCATION: 4276..4278
; FEATURE:
; NAME/KEY: polyadenylation site
; LOCATION: 4374..4379
US-09-486-580A-1

Alignment Scores:
Pred. No.: 3.9e-21      Length: 4415
Score: 209.00           Matches: 39
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 89.70%             Indels: 0
DB: 4                     Gaps: 0

US-10-045-180A-5 (1-44) x US-09-486-580A-1 (1-4415)

QY 1 GlnProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAspAsp 20
Db 3463 GAGCCGCTCCAGGCAAGAGCTCATGAGATGCCAGCCAGAGAGCCTCCAGAGATGAC 3522

QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 39
Db 3523 CAGATGTGCTCATTTACTTTTCAGAGATGACAGCTGCTCTTCAGGTTCCAGGCT 3579

RESULT 3
US-08-158-189-6
; Sequence 6, Application US/08158189
; Patent No. 5641497
; GENERAL INFORMATION:
; APPLICANT: Beying, Charles L.
; APPLICANT: Jones, Douglas E.
; TITLE OF INVENTION: Gastrointestinal Defensin Peptides,
; TITLE OF INVENTION: CDNA Sequences, Methods for Production and Use Thereof
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz and No. 5641497ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/158,189
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/888,232
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Johnson, Philip S.
; REGISTRATION NUMBER: 27,200
; REFERENCE/DOCKET NUMBER: CH-0219
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; FEATURE:
; NAME/KEY: CDS
```

LOCATION: 19..321
US-08-158-189-6

Alignment Scores:

Pred. No.:	2,38e-06	Length:	452
Score:	101.00	Matches:	21
Percent Similarity:	71.43%	Conservative:	9
Best Local Similarity:	50.00%	Mismatches:	12
Query Match:	43.35%	Indels:	0
DB:	1	Gaps:	0

US-10-045-180A-5 (1-44) x US-08-158-189-6 (1-452)

QY 1 GUPProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
DB 97 GATCCACTGCAGCAAAAGCTTATGAGCTGATGCCAGAGACGCTGGGCAATATGAC 156
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 157 CAGGACTTGGCGCTCTCCCTTTCAGAGAGATGCAAGCTCAAGTCTTAGAGCTTTGGGCTCA 216

QY 41 ThrIys 42
DB 217 ACAAGG 222

RESULT 4

US-09-016-434-1198
Sequence 1198, Application US/09016434

GENERAL INFORMATION:

APPLICANT: Janice Au-Young
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1198:
SEQUENCE CHARACTERISTICS:
LENGTH: 452 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G181546
US-09-016-434-1198

Alignment Scores:

Pred. No.:	2,38e-06	Length:	452
Score:	101.00	Matches:	21
Percent Similarity:	71.43%	Conservative:	9
Best Local Similarity:	50.00%	Mismatches:	12
Query Match:	43.35%	Indels:	0
DB:	4	Gaps:	0

US-10-045-180A-5 (1-44) x US-09-016-434-1198 (1-452)

QY 1 GUPProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
DB 97 GATCCACTGCAGCAAAAGCTTATGAGCTGATGCCAGAGACGCTGGGCAATATGAC 156
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 157 CAGGACTTGGCGCTCTCCCTTTCAGAGAGATGCAAGCTCAAGTCTTAGAGCTTTGGGCTCA 216
QY 41 ThrIys 42
DB 217 ACAAGG 222

RESULT 5

US-09-309-487-28
Sequence 28, Application US/09309487

GENERAL INFORMATION:

APPLICANT: Selsted, Michael E.
APPLICANT: Tang, Yi-Quan
APPLICANT: Yuan, Jun
APPLICANT: Chellette, Andre J.
TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
FILE REFERENCE: P-UC 3095
CURRENT APPLICATION NUMBER: US/09/309,487
CURRENT FILING DATE: 1999-05-10
NUMBER OF SEQ ID NOS: 31
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 28
LENGTH: 243
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (7)..(237)
US-09-309-487-28

Alignment Scores:

Pred. No.:	0.000184	Length:	243
Score:	86.00	Matches:	20
Percent Similarity:	61.36%	Conservative:	7
Best Local Similarity:	45.45%	Mismatches:	17
Query Match:	36.91%	Indels:	0
DB:	4	Gaps:	0

US-10-045-180A-5 (1-44) x US-09-309-487-28 (1-243)

QY 1 GUPProLeuGlnAlaArgAlaHisGluMetProAlaGlnIlyGlnProProAlaAsp 20
DB 67 GAGGACGCTGCAGCAAAAGCTTATGAGCTGATGCCAGAGACGCTGGGCAATATGAC 126
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
DB 127 CAGGGAATGGCTCATCTTCAATGAGCTGAAAGCGCGCTCTTCCACTTCAGAGTCA 186
QY 41 ThrIysGlyLeu 44
DB 187 GCGAAAGGCTTG 198

RESULT 6

US-09-967-808-28
Sequence 28, Application US/09967808
Patent No. 6514727
GENERAL INFORMATION:

```

; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; TITLE OF INVENTION: Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 28
; LENGTH: 243
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (7)..(237)
; US-09-967-808-28

```

Alignment Scores:

```

Pred. No.: 0.000184 Length: 243
Score: 86.00 Matches: 20
Percent Similarity: 61.36% Conservative: 7
Best Local Similarity: 45.45% Mismatches: 17
Query Match: 36.91% Indels: 0
DB: 4 Gaps: 0

```

US-10-045-180a-5 (1-44) x US-09-967-808-28 (1-243)

```

QY 1 GIuProlEuglInAlaArgAlaHisgluMetProAlaGlnLySglInProProAlaasp 20
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 GAGGCAAGCTGACGAGAGAGCTGATGAGCTGCCGCCAGACAGCTTGAGAGCATGAT 126
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 CAGGGAATGGCTCATTCCTTTACATGGCTGAAAGCCGCTCTTCACCTTTCAGAGTCA 186
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 41 ThrLySgLYleu 44
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 GCGAAGAGCTTG 198
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 7

```

; US-09-309-487-15
; Sequence 15, Application US/09309487
; Patent No. 6335318
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using Same
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/309,487
; CURRENT FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 15
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Macaca mulatta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (90)..(320)
; US-09-309-487-15

```

Alignment Scores:

```

Pred. No.: 0.00125 Length: 495
Score: 83.00 Matches: 19
Percent Similarity: 61.36% Conservative: 8
Best Local Similarity: 43.18% Mismatches: 17

```

```

Query Match: 35.62% Indels: 0
DB: 4 Gaps: 0

```

US-10-045-180a-5 (1-44) x US-09-309-487-15 (1-495)

```

QY 1 GIuProlEuglInAlaArgAlaHisgluMetProAlaGlnLySglInProProAlaasp 20
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 GAGGCAAGCTGACGAGAGAGCTGATGAGCTGCCGCCAGACAGCTTGAGAGCATGAT 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 CAGGGAATGGCTCATTCCTTTACATGGCTGAAAGCCGCTCTTCGCTTTCAGAGTCA 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 41 ThrLySgLYleu 44
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 GCGAAGAGCTTG 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8

```

; US-09-967-808-15
; Sequence 15, Application US/09967808
; Patent No. 6514727
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan
; APPLICANT: Yuan, Jun
; APPLICANT: Ouellette, Andre J.
; TITLE OF INVENTION: Antimicrobial Theta Defensins and Methods of Using
; FILE REFERENCE: P-UC 3095
; CURRENT APPLICATION NUMBER: US/09/967,808
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/309,487
; PRIOR FILING DATE: 1999-05-10
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 15
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Macaca mulatta
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (90)..(320)
; US-09-967-808-15

```

Alignment Scores:

```

Pred. No.: 0.00125 Length: 495
Score: 83.00 Matches: 19
Percent Similarity: 61.36% Conservative: 8
Best Local Similarity: 43.18% Mismatches: 17
Query Match: 35.62% Indels: 0
DB: 4 Gaps: 0

```

US-10-045-180a-5 (1-44) x US-09-967-808-15 (1-495)

```

QY 1 GIuProlEuglInAlaArgAlaHisgluMetProAlaGlnLySglInProProAlaasp 20
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 GAGGCAAGCTGACGAGAGAGCTGATGAGCTGCCGCCAGACAGCTTGAGAGCATGAT 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 21 GlnAspValValIleTyrPheSerGlyAspAspSerCysSerLeuGlnValProGlySer 40
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 210 CAGGGAATGGCTCATTCCTTTACATGGCTGAAAGCCGCTCTTCGCTTTCAGAGTCA 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 41 ThrLySgLYleu 44
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 270 GCGAAGAGCTTG 281
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 9

```

; US-09-309-487-13
; Sequence 13, Application US/09309487
; Patent No. 6335318
; GENERAL INFORMATION:
; APPLICANT: Selsted, Michael E.
; APPLICANT: Tang, Yi-Quan

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 886.963 Seconds
(without alignments)
1429.821 Million cell updates/sec

Title: US-10-045-180a-6

Perfect score: 191
Sequence: 1 ICHCRVLYCIPGHHGTCFLIGERYPICCY 31

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 segs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlp
-Q/cgnt2.1/USPTO_spool_p/US10045180/runat.17122003.145023.12528/app.query.fasta.1.860
-DB=GenEmbl -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-DOCALLS-bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human40.cdi -DIST=45
-OUTFMT=pcio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180 @CGN 1.1 5283 @runat.17122003.145023.12528 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEJUTERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGIOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl:*
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_ov:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	191	100.0	453	6	A98571	A98571 Sequence 2
2	191	100.0	453	6	BD074746	BD074746 Human def
3	191	100.0	4415	6	A98570	A98570 Sequence 1
4	191	100.0	4415	6	BD074745	BD074745 Human def
5	191	100.0	128544	9	AF238378	AF238378 Homo sapi
6	159	83.2	163093	2	AC116559	AC116559 Papio anu
7	159	83.2	188604	2	AC116558	AC116558 Papio anu
8	101	52.9	271979	2	AC098442	AC098442 Rattus no
9	101	52.9	301130	2	AC128185	AC128185 Rattus no
10	89	46.6	347	9	AF188268	AF188268 Macaca mu
11	89	46.6	435	9	AF184160	AF184160 Macaca mu
12	87	45.5	433	9	AF184159	AF184159 Macaca mu
13	85.5	44.8	1260	10	RN050356	RN050356 Rattus norv
14	84	44.0	494	10	RN016684	RN016684 Rattus norv
15	84	44.0	504	10	RN016683	RN016683 Rattus norv
16	84	44.0	3003	10	RN050354	RN050354 Rattus norv
17	84	44.0	3046	10	RN050353	RN050353 Rattus norv
18	84	44.0	3400	10	RN050355	RN050355 Rattus norv
19	84	44.0	242829	2	AC114391	AC114391 Rattus no
20	82	42.9	67979	2	AC113099	AC113099 Mus muscu
21	81	42.4	347	9	AF188270	AF188270 Macaca mu
22	81	42.4	485	4	RABDNP4A	M64601 Oryctolagus
23	80	41.9	90	6	BD128611	BD128611 Method fo
24	80	41.9	108	6	I32861	I32861 Sequence 43
25	80	41.9	108	6	I32862	I32862 Sequence 44
26	80	41.9	108	6	I32863	I32863 Sequence 45
27	80	41.9	109	6	I32878	I32878 Sequence 61
28	80	41.9	316	9	HUMDEF3A	M21131 Homo sapien
29	80	41.9	448	9	HUMDEF1A	M21130 Human neutr
30	80	41.9	451	9	HUMDEF2A	M23281 Human defen
31	80	41.9	464	9	HSDEF3	X13621 Human HNP-3
32	80	41.9	478	6	BD028213	BD028213 Sequence
33	80	41.9	482	9	HSHP1	X52053 Human mRNA
34	80	41.9	498	6	BD174049	BD174049 Method of
35	80	41.9	498	9	HUMDEF1AA	M26602 Human defen
36	80	41.9	509	6	BD023843	BD023843 Sequence
37	80	41.9	514	6	AX405718	AX405718 Sequence
38	80	41.9	543	6	AX588765	AX588765 Sequence
39	80	41.9	556	9	BC027917	BC027917 Homo sapi
40	80	41.9	1097	6	I30337	I30337 Sequence 5
41	80	41.9	1097	6	I96061	I96061 Sequence 5
42	80	41.9	2147	17	AF119902	AF119902 Homo sapi
43	80	41.9	3710	6	BD174050	BD174050 Method of
44	80	41.9	3710	9	HUMNTRT1	L12690 Human neut
45	80	41.9	3710	9	HUMNTRT1I	L12691 Human neut

RESULT 1

ALIGNMENTS

A98571
LOCUS A98571 453 bp DNA linear PAT 26-JAN-2000
DEFINITION Sequence 2 from Patent WO9911663.
ACCESSION A98571
VERSION A98571.1 GI:6781627
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Bougueleret, L., and Chumakov, I.
TITLE HUMAN DEFENSIN DEF-X, GENE AND DNAC, COMPOSITION CONTAINING SAME
JOURNAL AND DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
Bougueleret, L., and Chumakov, I.
Patent: WO 9911663-A 2 11-MAR-1999;
BOUGUELERET LYDIE (FR); CHUMAKOV ILVA (FR)
FEATURES
source location/Qualifiers
1..453
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
BASE COUNT 107 a 121 c 97 g 128 t
ORIGIN
Alignment Scores:
Pred. No.: 5.7e-19 Length: 453
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-045-180A-6 (1-31) X A98571 (1-453)
QY 1 11leCysHsCysArgValLeuTyrCys11lePheGlyGluHisLeuGlyGlyThrCysPhe 20
Db 241 ATCTGCATTCAGAGAGTACTACTGCACTTTTGGAGAACATCTTGGAGACCTGCTTC 300
QY 21 11leuGlyGluArgTyrPro11leCysCysTyr 31
Db 301 ATCCTTGCTGAACGCTACCAATCTGCTGCTAC 333
RESULT 2
BD074746 453 bp DNA linear PAT 27-AUG-2002
LOCUS BD074746
DEFINITION Human defensin polypeptide Def-X, genome DNA and cDNA, composition
containing the same, and application to diagnosis and remedy
treatment.
ACCESSION BD074746
VERSION BD074746.1 GI:22620349
KEYWORDS
SOURCE JP 2001514264-A/2.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
AUTHORS Bougueleret, L., and Shmacov, I.
TITLE Human defensin polypeptide Def-X, genome DNA and cDNA, composition
containing the same, and application to diagnosis and remedy
JOURNAL Patent: JP 2001514264-A 2 11-SEP-2001;
GENSET
OS Homo sapiens (human)
PN JP 2001514264-A/2
PD 11-SEP-2001
PF 28-AUG-1998 JP 2000508701
PR 29-AUG-1997 FR 97/10823
PI LYDIE BOUGUELERET, ILVA SHMACOV
PC C07K14/435, A01N43/50, A01N63/00, A61K7/00, A61K38/00, A61P29/00,
PC A61P35/00,
PC A61P37/02, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N15/09, PC
C12P21/02,
PC C12P21/08, C12Q1/68, G01N33/53, A61K37/02, C12N15/00 CC
Strandedness: Double;
CC Topology: linear;

CC Human defensin polypeptide Def-X, genome DNA and cDNA, CC
composition
containing the same, and application to diagnosis and remedy
treatment
CC CC
FH Key location/Qualifiers
FT source 1..453
/organism="Homo sapiens (human)".
FEATURES
source location/Qualifiers
1..453
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 107 a 121 c 97 g 128 t
ORIGIN
Alignment Scores:
Pred. No.: 5.7e-19 Length: 453
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-045-180A-6 (1-31) X BD074746 (1-453)
QY 1 11leCysHsCysArgValLeuTyrCys11lePheGlyGluHisLeuGlyGlyThrCysPhe 20
Db 241 ATCTGCATTCAGAGAGTACTACTGCACTTTTGGAGAACATCTTGGAGACCTGCTTC 300
QY 21 11leuGlyGluArgTyrPro11leCysCysTyr 31
Db 301 ATCCTTGCTGAACGCTACCAATCTGCTGCTAC 333
RESULT 3
A98570 4415 bp DNA linear PAT 26-JAN-2000
LOCUS A98570
DEFINITION Sequence 1 from Patent WO9911663.
ACCESSION A98570
VERSION A98570.1 GI:6781626
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Bougueleret, L., and Chumakov, I.
TITLE HUMAN DEFENSIN DEF-X, GENE AND DNAC, COMPOSITION CONTAINING SAME
AND DIAGNOSTIC AND THERAPEUTIC APPLICATIONS
JOURNAL Patent: WO 9911663-A 1 11-MAR-1999;
BOUGUELERET LYDIE (FR); CHUMAKOV ILVA (FR)
FEATURES
source location/Qualifiers
1..4415
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
1836..1874
exon 3394..3577
exon 4161..4380
exon 4374..4379
BASE COUNT 1128 a 1109 c 912 g 1252 t 14 others
ORIGIN
Alignment Scores:
Pred. No.: 5.46e-18 Length: 4415
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-045-180A-6 (1-31) X A98570 (1-4415)
QY 1 11leCysHsCysArgValLeuTyrCys11lePheGlyGluHisLeuGlyGlyThrCysPhe 20
|||||

Db 4181 ATCTGCATTCGACGACTACTACTGATTTTGGAGAACATCTTGTTGGAGACTGCTTC 4240

QY 21 |||leuugl|g|u|a|g|y|p|p|o|l|e|c|y|s|e|t|y|r 31
 |||leuugl|g|u|a|g|y|p|p|o|l|e|c|y|s|e|t|y|r 31

Db 4241 ATCTTGTTGTAACGCTACCACTCTGCTGCTAC 4273

RESULT 4

BD074745 4415 bp DNA linear PAT 27-AUG-2002

LOCUS Human defensin polypeptide Def-X, genome DNA and cDNA, composition

DEFINITION containing the same, and application to diagnosis and remedy treatment.

ACCESSION BD074745.1 GI:22620348

VERSION BD074745.1

KEYWORDS JP 2001514264-A/1.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Bougueleret, L. and Shmucov, I.

TITLE Human defensin polypeptide Def-X, genome DNA and cDNA, composition containing the same, and application to diagnosis and remedy

JOURNAL Patent: JP 2001514264-A 1 11-SEP-2001;

COMMENT GENSET

OS Homo sapiens (human)

PN JP 2001514264-A/1

PD 11-SEP-2001

PR 28-AUG-1998 JP 2000508701

PR 29-AUG-1997 FR 97/10823

P1 LYDIE BOUGUELERET, ILYA SHMUCOV

PC C07K14/435, A01N43/50, A01N63/00, A61K7/00, A61K38/00, A61P29/00, A61P35/00,

PC A61P37/02, C07K16/18, C12N1/15, C12N1/19, C12N1/21, C12N15/09, PC C12P21/02,

PC C12P21/08, C12Q1/68, G01N33/53, A61K37/02, C12N15/00 CC

Strandedness: Double;

CC Topology: Linear;

CC Human defensin polypeptide Def-X, genome DNA and cDNA, CC composition

CC containing the same, and application to diagnosis and remedy

CC treatment

CC Key Location/Qualifiers

CC FT exon 1836..1874

CC FT exon 3394..3577

CC FT exon 4161..4380

CC CDS 3406..3408

CC FT 4276..4278.

FEATURES

source

1..4415 Location/Qualifiers

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 1128 a 1109 c 912 g 1252 t 14 others

ORIGIN

Alignment Scores:

Pred. No.: 5.46e-18 Length: 4415

Score: 191.00 Matches: 31

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-10-045-180a-6 (1-31) x BD074745 (1-4415)

QY 1 |||leCyHisCySaryValleuYrCySilePheGlyGluHisleuGlyGlyThrCySph 20
 |||leCyHisCySaryValleuYrCySilePheGlyGluHisleuGlyGlyThrCySph 20

Db 4181 ATCTGCATTCGACGACTACTACTGATTTTGGAGAACATCTTGTTGGAGACTGCTTC 4240

QY 21 |||leuugl|g|u|a|g|y|p|p|o|l|e|c|y|s|e|t|y|r 31
 |||leuugl|g|u|a|g|y|p|p|o|l|e|c|y|s|e|t|y|r 31

Db 4241 ATCTTGTTGTAACGCTACCACTCTGCTGCTAC 4273

RESULT 5

AF238378/c 128544 bp DNA linear PRI 02-APR-2003

LOCUS Homo sapiens chromosome 8 clone SCB-561b17 map p22-p21, complete

DEFINITION sequence.

ACCESSION AF238378.5 GI:29469504

VERSION AF238378

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Polley, A., Baumgart, C., Blechschmidt, K., Detle, M.D., Jahn, N., Menzel, U., Reichwald, K., Schlinhabel, M.B., Schudy, A., Taudien, S., Wen, G., Schutte, B., Malik, M., Peng, J., Hong, M., McCray, P., and Rosenthal, A.

TITLE Chromosome 8 genomic sequence

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 128544)

AUTHORS Schudy, A., Schlinhabel, M., Schutte, B., Ganz, T., Linzmeier, R., Ho, C.H., Hoang, B.V., McCray, P., Baumgart, C., Menzel, U., Schattevoy, R., and Rosenthal, A.

TITLE Direct Submission

JOURNAL Submitted (22-FEB-2000) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

REFERENCE 3 (bases 1 to 128544)

AUTHORS Genome Sequencing Center Jena.

TITLE Direct Submission

JOURNAL Submitted (24-MAY-2000) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

REFERENCE 4 (bases 1 to 128544)

AUTHORS Genome Sequencing Center Jena.

TITLE Direct Submission

JOURNAL Submitted (10-NOV-2000) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

REFERENCE 5 (bases 1 to 128544)

AUTHORS Lagemann, D. and Platzner, M.

TITLE Direct Submission

JOURNAL Submitted (06-JUL-2002) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

REFERENCE 6 (bases 1 to 128544)

AUTHORS Lagemann, D. and Platzner, M.

TITLE Direct Submission

JOURNAL Submitted (02-APR-2003) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

COMMENT On Apr 2, 2003 this sequence version replaced GI:21700555.

Center: Institute of Molecular Biotechnology

Center code: IMB

Web site: http://genome.imb-jena.de/

Contact: gscj-submit@genome.imb-jena.de

Project Information

Center project name: H370

Center clone name: SCB-561b17

Summary Statistics

Sequencing vector: pUC18; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 127999 bases at least Q40

Consensus quality: 128330 bases at least Q40

Consensus quality: 128544 bases at least Q20

Quality coverage: 11.05x

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

```
DB      62490 ATCTGTGTAAGCTACCAATCTGCTGTAC 62458
RESULT 6
AC116559
LOCUS
DEFINITION
Papio anubis clone rp41-339c10, WORKING DRAFT SEQUENCE, 3 ordered
pieces.
AC116559
AC116559 163093 bp DNA linear HTG 30-MAY-2003
VERSION
AC116559.18 GI:31193968
KEYWORDS
HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE
Papio anubis (olive baboon)
ORGANISM
Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
1 (bases 1 to 163093)
XU,W., HUA,A., EICHLER,E. and ROE,B.A.
JOURNAL
Unpublished
2 (bases 1 to 163093)
XU,W., HUA,A., EICHLER,E. and ROE,B.A.
JOURNAL
Direct Submission
Submitted (29-MAR-2002) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 163093)
XU,W., HUA,A., EICHLER,E. and ROE,B.A.
JOURNAL
Direct Submission
Submitted (30-MAY-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 30, 2003 this sequence version replaced gi:30725962.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 6315: contig of 6315 bp in length
* 6316 6415: gap of unknown length
* 6416 16704: contig of 10289 bp in length
* 16705 16804: gap of unknown length
* 16805 163093: contig of 146289 bp in length.
Location/Qualifiers
1. 163093
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="rp41-339c10"
/clone_lib="RPCT - 41 Male (Olive) Baboon BAC Library"
BASE COUNT 44435 a 36508 c 35861 g 46065 t 224 others
ORIGIN
Alignment Scores:
Pred. No.: 1 08e-11 Length: 163093
Score: 159.00 Matches: 26
Percent Similarity: 90.00% Conservaive: 1
Best Local Similarity: 86.67% Mismatches: 3
Query Match: 83.25% Indels: 0
DB: 2 Gaps: 0
US-10-045-180a-6 (1-31) x AC116559 (1-163093)
QY 2 CysHsCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21

DB      33641 TCCCATGTGAGAGTACTGCTCTTTTGAGAAACATCTGTGGAGACTCTTCATC 33700
QY      22 LeuGlyGluArgTyrProIleCysCysTyr 31
DB      33701 CATGTGAAAGCTCCCAATCTGCTGTAC 33730
RESULT 7
AC116558
LOCUS
DEFINITION
Papio anubis clone rp41-273g19, WORKING DRAFT SEQUENCE, 2 ordered
pieces.
AC116558
AC116558 188604 bp DNA linear HTG 11-JUN-2003
VERSION
AC116558.16 GI:30725961
KEYWORDS
HTG; HTGS PHASE2; HTGS DRAFT.
SOURCE
Papio anubis (olive baboon)
ORGANISM
Papio anubis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Papio.
1 (bases 1 to 188604)
XU,W., HUA,A., EICHLER,E. and ROE,B.A.
JOURNAL
Unpublished
2 (bases 1 to 188604)
XU,W., HUA,A., EICHLER,E. and ROE,B.A.
JOURNAL
Direct Submission
Submitted (11-JUN-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
3 (bases 1 to 188604)
XU,W., HUA,A., EICHLER,E. and ROE,B.A.
JOURNAL
Direct Submission
Submitted (11-JUN-2003) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
On May 15, 2003 this sequence version replaced gi:28173120.
----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code:UOKNOR
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 6057: contig of 6057 bp in length
* 6058 6157: gap of unknown length
* 6158 188604: contig of 182447 bp in length.
Location/Qualifiers
1. 188604
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="rp41-273g19"
/clone_lib="RPCT - 41 Male (Olive) Baboon BAC Library"
BASE COUNT 53384 a 41769 c 40796 g 52555 t 100 others
ORIGIN
Alignment Scores:
Pred. No.: 1 25e-11 Length: 188604
Score: 159.00 Matches: 26
Percent Similarity: 90.00% Conservaive: 1
Best Local Similarity: 86.67% Mismatches: 3
Query Match: 83.25% Indels: 0
DB: 2 Gaps: 0
US-10-045-180a-6 (1-31) x AC116558 (1-188604)
```


BAC COUNT 74463 a 48993 c 50388 g 70500 t 27635 others
 ALIGNMENT SCORES:
 Pred. No.: 0 00702 Length: 271979
 Score: 101.00 Matches: 16
 Percent Similarity: 70.00% Conservative: 5
 Best Local Similarity: 53.33% Mismatches: 9
 Query Match: 52.88% Indels: 0
 DB: 2 Gaps: 0
 US-10-045-180a-6 (1-31) x AC098442 (1-271979)
 QY 2 CysHicCysArgValIleuTyrcysIlePheGlyGlyHicIleuGlyGlyThrCysPheIle 21
 Db 228016 TGTCTATTGCAGATCTCTCTCACTCTCACTTTGGGGAGACGACTCCCTGGCTCTCCCTCT 227957
 QY 22 LeuGlyGluArgTyrcysProIleCysCysTyrcys 31
 Db 227956 GGTGGCGTCACCTACCCCTCTGCTGCAC 227927
 RESULT 9
 AC128185 301130 bp DNA linear HTG 22-SEP-2002
 LOCUS Rattus norvegicus clone CH230-30N20, *** SEQUENCING IN PROGRESS
 DEFINITION *** 4 unordered pieces.
 AC128185 GI:23267661
 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 1 (bases 1 to 301130)
 Muzny,D.,Marle,M.,Metzker,M.,Lee,A.,Abramzon,S.,Adams,C.,Alder,J.,Allen,C.,Allen,H.,Alsbrooks,S.,Amin,A.,Anguiano,D.,Anyalebechi,V.,Aoyagi,A.,Ayodeji,M.,Bacca,R.,Baden,H.,Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,Biswal,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,Bryant,N.,Buhay,C.,Burck,P.,Buttelli,K.,Calderson,E.,Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Chen,A.,Chaco,J.,Chavez,D.,Chen,Y.,Chen,R.,Chen,Y.,Chen,Y.,Chu,J.,Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,David,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dederich,D.,Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Diya,K.,Draper,H.,Dugan-Rocha,S.,Dunn,A.,Dudin,K.,Duvall,B.,Eaves,K.,Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,Fraser,C.M.,Gabisi,A.,Ganta,R.,Garcia,A.,Garner,T.,Garza,M.,Gharageorgis,E.,Geer,K.,Gill,R.,Grady,M.,Guerra,W.,Guevara,W.,Gubaratne,P.,Haaland,W.,Hamil,C.,Hamilton,C.,Hamilton,K.,Harvey,Y.,Havlak,P.,Hawes,A.,Henderson,N.,Hernandez,J.,Hollins,B.,Hines,S.,Hladus,S.L.,Hodgson,A.,Hogues,M.,Hollins,B.,Howells,S.,Huliy,S.,Hume,J.,Idlebird,D.,Jackson,A.,Jackson,L.,Jacob,L.,Jiang,S.,Khan,Z.,King,L.,Koyar,C.,Karpethy,S.,Kelly,S.,Kelly,S.,Johnson,B.,Johnson,R.,Jolivet,A.,Kowis,C.,Kraft,C.L.,Ledow,H.,Levan,J.,Lewis,L.,Li,Z.,Lin,J.,Lin,J.,Lin,W.,Lin,Y.,London,P.,Longacre,S.,Lopez,D.,Lorensuewa,L.,Louisege,H.,Lozado,R.J.,Lu,X.,Ma,J.,Maheshwari,M.,Mahindartine,M.,Mahmoud,M.,Malloy,K.,Mangum,A.,Mangum,B.,Mapua,P.,Martin,K.,Martin,R.,Martinez,E.,Mawhney,S.,McLeod,M.P.,McNeill,T.Z.,Meenen,B.,Millojaevljic,A.,Miner,G.,Minga,E.,Montemayor,J.,Moore,S.,Morgan,M.,Morris,K.,Morris,S.,Mundasa,M.,Murphy,M.,Nair,L.,Nankervis,C.,Neal,D.,Newton,G.,Olanrunpaocoon,A.,Palis,S.,Naokeleneh,O.,Okunou,G.,Olanrunpaocoon,A.,Palis,S.,Parks,K.,Pasternak,S.,Paul,H.,Perez,A.,Perez,L.,Pfankoch,C.,Plopper,F.,Poindexter,A.,Popovic,D.,Primus,E.,Pu,L.L.,Puzo,M.,Quiroz,J.,Rachin,E.,Reeves,K.,Regier,M.A.,Reigh,R.,Reilly,B.,Reilly,M.,Ren,Y.,Reuter,M.,Richards,S.,Riggs,F.,Rives,C.,Rockey,T.,Rojas,A.,Rose,M.,Richards,S.,Ruiz,S.J.,Sanders,W.,Savery,G.,Scherer,S.,Scott,G.,Shatsman,S.,Shen,H.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GYND
 Center clone name: CH230-30N20
 Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 240138 bases at least Q40
 Consensus quality: 244747 bases at least Q30
 Consensus quality: 247371 bases at least Q20
 Estimated insert size: 272390; sum-of-ctnigs estimation
 Quality coverage: 3x in Q20 bases; sum-of-ctnigs estimation
 NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 NOTE: This sequence may represent more than one clone.
 NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 1 3552: contig of 3552 bp in length
 3553 3652: gap of unknown length
 3653 16658: contig of 13006 bp in length
 16659 16758: gap of unknown length
 16759 296490: contig of 279732 bp in length
 296491 296590: gap of unknown length
 296591 301130: contig of 4540 bp in length.
 Location/Qualifiers
 1. 301130
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-30N20"

QY 22 LeuGIyGIuargTyrProIleCysCys 30

Accession	Definition
U16604	Rattus norvegicus defensin RatNP-4 precursor mRNA, complete cds.

VERSION U16684.1 GI:1041806
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 494)
AUTHORS Yount,N.Y., Wang,M.S., Yuan,J., Banaiee,N., Ouellette,A.J. and
Selssted,M.E.
TITLE Rat neutrophil defensins. Precursor structures and expression
during neutrophilic myelopoiesis
JOURNAL J. Immunol. 155 (9), 4476-4484 (1995)
MEDLINE 96025910
PUBMED 7594610
REFERENCE 2 (bases 1 to 494)
AUTHORS Yount,N.Y.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-1994) Nannette Y. Yount, Pathology, University of
California, D440 Med. Sci. I, Irvine, CA 92717, USA
FEATURES
source
1..494
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="bone marrow"
/clone_lib="plasmid, oligo dt primed"
72..353
/codon_start=1
/product="defensin RatNP-4 precursor"
/protein_id="AA91972.1"
/db_xref="GI:1041807"
/translation="MRTLTTLTLALHTQAESPOERAKAPDDQVMEDQDFIS
FGYKGVLDAAVKAQACYCIRGACVSGERLTGACGLNGRIYRLCC"
CDS
sig_peptide
mat_peptide
repeat_region
polya_signal
polya_site
BASE COUNT 123 a 119 c 129 g 123 t
ORIGIN
Alignment Scores:
Pred. No.: 0.00442 Length: 494
Score: 84.00 Matches: 13
Percent Similarity: 62.07% Conservative: 5
Best Local Similarity: 44.83% Mismatches: 11
Query Match: 43.98% Indels: 0
DB: 10 Gaps: 0
US-10-045-180A-6 (1-31) x RNU16684 (1-494)
QY 2 CysHsCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
Db 261 TGCTATTGCAAGATCGAGCCTGTGTCTTCTGAGAAAGCGCTCAGCGGGCATGTGCTC 320
QY 22 LeuGlyGluArgTyrProIleCysCys 30
Db 321 AATGGCCGCACTACCGCTCTGTGTC 347
RESULT 15
RNU16683 504 bp mRNA linear ROD 12-MAR-1996
LOCUS Rattus norvegicus defensin RatNP-3 precursor mRNA, complete cds.
DEFINITION
ACCESSION U16683
VERSION U16683.1 GI:1041804
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 504)
AUTHORS Yount,N.Y., Wang,M.S., Yuan,J., Banaiee,N., Ouellette,A.J. and
Selssted,M.E.
TITLE Rat neutrophil defensins. Precursor structures and expression
during neutrophilic myelopoiesis
JOURNAL J. Immunol. 155 (9), 4476-4484 (1995)
MEDLINE 96025910
PUBMED 7594610
REFERENCE 2 (bases 1 to 504)
AUTHORS Yount,N.Y.
TITLE Direct Submission
JOURNAL Submitted (27-OCT-1994) Nannette Y. Yount, Pathology, University of
California, D440 Med. Sci. I, Irvine, CA 92717, USA
FEATURES
source
1..504
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/tissue_type="bone marrow"
/clone_lib="plasmid, oligo dt primed"
62..325
/codon_start=1
/product="defensin RatNP-3 precursor"
/protein_id="AA91971.1"
/db_xref="GI:1041805"
/translation="MRTLTTLTLALHTQAESPOGSKAPDEBDISVFGDGK
GTALDDAAVKAQVTCSTSSCRFSERLSCAGRLNGRIYRLCC"
CDS
sig_peptide
mat_peptide
repeat_region
polya_signal
polya_site
BASE COUNT 149 a 117 c 121 g 117 t
ORIGIN
Alignment Scores:
Pred. No.: 0.0045 Length: 504
Score: 84.00 Matches: 14
Percent Similarity: 55.17% Conservative: 2
Best Local Similarity: 48.28% Mismatches: 13
Query Match: 43.98% Indels: 0
DB: 10 Gaps: 0
US-10-045-180A-6 (1-31) x RNU16683 (1-504)
QY 2 CysHsCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
Db 226 TGCTTTTAAAGCTCAAGCTGTGCTTTTGGAGAAAGGCTCTCGGGCATGTGCTC 295
QY 22 LeuGlyGluArgTyrProIleCysCys 30
Db 296 AATGGCCGCACTACCGACTCTGTGTC 322
Search completed: December 17, 2003, 16:40:23
Job time : 961.963 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 1083.85 Seconds
(without alignments)
695.153 Million cell updates/sec

SUMMARIES

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Title: US-10-045-180a-6
Perfect score: 191
Sequence: 1 ICHCRVLYCIFGSEHTGTCFTIGHERPICCY 31

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delep 6.0, Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-Processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlp
-=/cgn2.1/USPTO.spool.p/US10045180/rnmat.17122003.145024.12540/app_query.fasta.1.860
-DB=EST -OPMT=fast3d -SUFFIX=rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cgi -LIST=45
-DOCALIGN=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=Pct -NORM=ext -HEAPSIZE=500 -MNLLEN=0 -MAXLEN=2000000000
-USER=US10045180@cgn.1.6100@rnmat.17122003.145024.12540 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEOUTERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -NARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estmu:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estom:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_pro:*
26: em_gss_rod:*
27: em_gss_vrl:*
28: gb_gss1:*

Result No.	Score	% Match	Query Length	DB ID	Description
C 1	164	85.9	606	28	A0624584
C 2	88	46.1	329	10	BF904561
C 3	84	44.0	454	9	A1408357
C 4	84	44.0	468	10	BF290038
C 5	84	44.0	547	13	BQ200749
C 6	82	42.9	287	9	AM467913
C 7	82	42.9	349	12	B1021835
C 8	80	41.9	181	13	BQ343473
C 9	80	41.9	219	10	BE831361
C 10	80	41.9	225	9	AA321199
C 11	80	41.9	229	9	AA321481
C 12	80	41.9	233	9	AM075767
C 13	80	41.9	234	10	BF93374
C 14	80	41.9	241	12	B1028909
C 15	80	41.9	246	10	BF903426
C 16	80	41.9	251	12	B1023343
C 17	80	41.9	255	10	BE831399
C 18	80	41.9	257	10	BE831404
C 19	80	41.9	259	9	AM274523
C 20	80	41.9	260	10	BE831426
C 21	80	41.9	269	12	B1022826
C 22	80	41.9	273	10	BE831401
C 23	80	41.9	274	10	BF901251
C 24	80	41.9	274	10	BF901407
C 25	80	41.9	274	10	BF901500
C 26	80	41.9	274	10	BF903483
C 27	80	41.9	274	10	BF905787
C 28	80	41.9	279	9	A1435021
C 29	80	41.9	286	13	BQ34909
C 30	80	41.9	288	10	BE831389
C 31	80	41.9	289	10	BE831385
C 32	80	41.9	290	10	BE831314
C 33	80	41.9	290	10	BE831369
C 34	80	41.9	290	10	BE831382
C 35	80	41.9	290	10	BE831396
C 36	80	41.9	290	10	BE831397
C 37	80	41.9	290	10	BE831398
C 38	80	41.9	290	10	BE831402
C 39	80	41.9	290	10	BE831405
C 40	80	41.9	290	10	BE831406
C 41	80	41.9	290	10	BE831407
C 42	80	41.9	290	10	BE831411
C 43	80	41.9	290	10	BE831414
C 44	80	41.9	291	10	BE831311
C 45	80	41.9	298	10	BE831364

ALIGNMENTS

RESULT 1
A0624584/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

A0624584
HS_2110 B1.D06 MR.CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2110 Col=11 Row=H, genomic survey sequence.
A0624584
GI:5086976
GSS.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

606 bp DNA linear GSS 16-JUN-1999

REFERENCE 1 (bases 1 to 606)
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
MEDLINE 99380589
PUBMED 10449764
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: <http://www.htsc.washington.edu>
Plate: 2110 row: H column: 11
Seq primer: M13 Reverse
Class: BAC ends
Location/Qualifiers
1..606
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="plate=2110 Col=11 Row=H"
/sex="male"
/clone_lib="CIT Approved Human Genomic Sperm Library D"
/note="Organ: Sperm; Vector: pBeloBAC11; BAC clones in E-Coli DH10B"

BASE COUNT 169 a 128 c 131 g 172 t 6 others
ORIGIN

Alignment Scores:
Pred. No.: 2.59e-11 Length: 606
Score: 164.00 Matches: 28
Percent Similarity: 90.32% Conservative: 0
Best Local Similarity: 90.32% Mismatches: 3
Query Match: 85.86% Indels: 0
DB: 28 Gaps: 0

US-10-045-180A-6 (1-31) x AG624584 (1-606)

QY 1 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe 20
Db 150 ATCTGCATTCGACAGACTACTACTGCACTTTTGGAGAACATCTTGGAGACAAATTC 91

QY 21 IleLeuGlyGluArgTyrProIleCysCysTyr 31
Db .90 ATCCTGTGTAACGCTACCCCATCTGCTCTAC 58

RESULT 2
BF904561/c 329 bp mRNA linear EST 18-JAN-2001
LOCUS PF3-MT0200-211200-002-all MT0200 Homo sapiens cDNA, mRNA sequence.
DEFINITION BF904561
VERSION BF904561.1 GI:12296020
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 329)
Dias Nero,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brumstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?ti=PF3&t2=PF3-MT0200-211200-002-all&t3=2000-12-21&t4=1>)
Seq primer: puc 18 forward
High quality sequence start: 2
High quality sequence stop: 329.
Location/Qualifiers
1..329
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0200"
/note="Organ: marrow; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 76 a 82 c 85 g 85 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 0.0852 Length: 329
Score: 88.00 Matches: 14
Percent Similarity: 58.62% Conservative: 3
Best Local Similarity: 48.28% Mismatches: 12
Query Match: 46.07% Indels: 0
DB: 10 Gaps: 0

US-10-045-180A-6 (1-31) x BF904561 (1-329)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe 21
Db 194 TGTATTGCAAGATACCAAGTGCATTCGACAGGAAAGCTGCTATGGAACCTGCATCTAC 135

QY 22 LeuGlyGluArgTyrProIleCysCys 30
Db 134 CAGGAAAGACTCTGCGCATTCCTCTGC 108

RESULT 3
A1408357/c 454 bp mRNA linear EST 09-FEB-1999
LOCUS A1408357
DEFINITION EST226647 Normalized rat ovary, Bento Soares Rattus sp. cDNA clone
VERSION ROVER82.3' end, mRNA sequence.
KEYWORDS A1408357 GI:4251861
SOURCE EST.
Rattus sp.
Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus sp.
1 (bases 1 to 454)
Lee,N.H., Glodde,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.
Rat Genome Project: Generation of a Rat EST (RST) Catalog & Rat Gene Index
Unpublished
Contact: Lee, NH

The Institute for Genomic Research
9712, Medical Center Drive, Rockville, MD 20850, USA
Tel: (301)-838-3529
Fax: (301)-838-0208
Email: nhlee@igr.org
Seq primer: M13-21.

FEATURES

source

Location/Qualifiers
1..454
/organism="Rattus sp."
/mol_type="mRNA"
/db_xref="taxon:10118"
/clone_1lb="ROVEP82"
/clone_1lb="Normalized rat ovary, Bento Soares"
/note="Organ: ovary; Vector: pT73Pac; Site_1: EcoRI;
Site_2: NotI"
BASE COUNT 111 a 118 c 106 g 119 t
ORIGIN

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
84.00	62.07%	454	13
Best Local Similarity:	44.83%	Conservative:	5
Query Match:	43.98%	Mismatches:	11
		Indels:	0
		Gaps:	0

US-10-045-180a-6 (1-31) x A1408357 (1-454)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyIuHisLeuGlyGlyThrCysPheIle 21

Db 239 TCCTATGCGAGATCGAGACCTGTGTCTGGAGAACGCTACTGGGCGATGTGCTTC 180

QY 22 LeuGlyGluArgTyrProIleCysCys 30

Db 179 AATGGCCGATCTACCGTCTGTGTC 153

RESULT 4

BP290038/c 468 bp mRNA linear EST 28-NOV-2000

LOCUS EST154629 Rat Gene Index, normalized rat, Rattus norvegicus cDNA

DEFINITION Rattus norvegicus cDNA clone RGIHM93 3' sequence, mRNA sequence.

ACCESSION BP290038

VERSION BP290038.1 GI:11221108

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 468)

Malek,R.L., Cho,J., Lee,Y., Karamycheva,S., Parvizl,B., Perrea,G.,

Sultana,R., Teal,J., White,J., Quackenbush,J. and Lee,N.H.

Generation of ESTs from Normalized Rat Embryo, Bento Soares

Unpublished

Other ESTs: EST352370

Contact: Lee, NH

The Institute for Genomic Research

9712, Medical Center Drive, Rockville, MD 20850, USA

Tel: (301)-838-3529

Fax: (301)-838-0208

Email: nhlee@igr.org

This clone is available through the ATCC, contact the ATCC

tel#703-365-2700 for further information.

Location/Qualifiers

1..468

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="RGIM93"

/tissue_type="mixed tissue"

/lab_host="DH5-alpha"

/clone_1lb="Rat Gene Index, normalized rat, Rattus

norvegicus cDNA"

BASE COUNT 117 a 122 c 110 g 119 t

ORIGIN

/note="Vector: pT73Pac; Site_1: EcoRI; Site_2: NotI;
Combination of KOV, RBR, RKL, RLI, RPL, RLU, REM, RRU, RSP
, RHE, RPC, RPN"

Alignment Scores:

Pred. No.:	Score:	Length:	Matches:
84.00	62.07%	468	13
Best Local Similarity:	44.83%	Conservative:	5
Query Match:	43.98%	Mismatches:	11
		Indels:	0
		Gaps:	0

US-10-045-180a-6 (1-31) x BP290038 (1-468)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyIuHisLeuGlyGlyThrCysPheIle 21

Db 239 TCCTATGCGAGATCGAGACCTGTGTCTGGAGAACGCTACTGGGCGATGTGCTTC 180

QY 22 LeuGlyGluArgTyrProIleCysCys 30

Db 179 AATGGCCGATCTACCGTCTGTGTC 153

RESULT 5

BQ200749/c 547 bp mRNA linear EST 02-MAY-2002

LOCUS UI-R-EB1-clg-e-09-0-UI.s1 UI-R-EB1 Rattus norvegicus cDNA clone

DEFINITION UI-R-EB1-clg-e-09-0-UI 3', mRNA sequence.

ACCESSION BQ200749

VERSION BQ200749.1 GI:20417214

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Rattus norvegicus

COMMENT Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 547)

Bonaldi,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

JOURNAL GENOME RESEARCH

PUBMED 97044477

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 9550

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

The sequence contained an oligo-dT track that was present in the

oligonucleotide that was used to prime the synthesis of first

strand cDNA and therefore this may represent a bonafide poly A

tail. The sequence tag present in the cDNA between the NotI site

and the oligo-dT track served to verify it as a clone from the

normalized duodenum library cDNA library Preparation: M.B. Soares

lab Clone distribution: clones will be available through Research

Genetics (www.reagen.com) The following repetitive elements were

found in this cDNA sequence: 118-200, >(GAA)n\$imple_repeat

Seq primer: M13 Forward

POLYA=Yes.

FEATURES

source

1..547

/organism="Rattus norvegicus"

/mol_type="mRNA"

/strain="Sprague-Dawley"

/db_xref="taxon:10116"

/clone="UI-R-EB1-clg-e-09-0-UI"

/dev_strage="adulr"

/lab_host="DH10B (Life Technologies)"

/clone_1lb="UI-R-EB1"

/note="Vector: pT73D-Pac (Pharmacia) with a modified

polylinker. Site_1: Not I; Site_2: Eco RI; UI-R-EB1 is a non-normalized Rat Duodenum library (RDU) constructed in pT37 PAC vector according to the procedure described by Bonaldo, Lennon & Soares (Genome Research Genome 6: 791-806, 1996). The oligonucleotide used to prime first strand synthesis contained the sequence tag TGCTGTTCA between the Not I cloning site and dT18 stretch. The Rat Duodenum tissue was provided by Tom Freeman of the Sanger Center.

BASE COUNT 120 a 130 c 119 g 178 t
TAG-TISSUE=duodenum
TAG_SEQ=TGCTGTTCA
ORIGIN

Alignment Scores:

Pred. No.: 0.557 Length: 547
Score: 84.00 Matches: 14
Percent Similarity: 55.17% Conservative: 2
Best Local Similarity: 48.28% Mismatches: 13
Query Match: 43.98% Indels: 0
DB: 13 Gaps: 0

US-10-045-180a-6 (1-31) x BQ200749 (1-547)

QY 2 CysHsCysArGValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
Db 295 TGCTCTGTAGAACCTCAAGCTGTCGTTGGAAAAGGCTCTCTGGGCGATGCGCTC 236
QY 22 LeuGlyGluArGlyTyrProIleCysCys 30
Db 235 AATGGCCGACATCTACCGCATCTGTTGC 209

RESULT 6
LOCUS AM467913 287 bp mRNA linear EST 24-FEB-2000
DEFINITION he28d03.x1 NCI CGAP CML1 Homo sapiens cDNA clone IMAGE:2920325 3'
Similar to gb:M6602 NEUTROPHIL DEFENSINS 1, 2 AND 3 PRECURSOR (HUMAN); mRNA sequence.
ACCESSION AM467913
VERSION AM467913.1 GI:7038019
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 287) NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Elisabeth Paietta, Jonathan D. Licht, M.D., Michael R. Emmert-Buck, M.D., Ph.D. CDNA Library Preparation: Life Technologies, Inc. CDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/INL at: www.bio.11nl.gov/db/tp/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 231.

FEATURES

source
1..287
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2920325"
/tissue_type="myeloid cells, 18 pooled CML cases, BCR/ABL rearrangement positive, includes both chronic phase and myeloid blast crisis"
/lab_host="DH10B"

/clone_1b="NCI CGAP CML1"
/note="Organ: whole blood; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies."
BASE COUNT 71 a 65 c 70 g 80 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 0.427 Length: 287
Score: 82.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 42.93% Indels: 0
DB: 9 Gaps: 0

US-10-045-180a-6 (1-31) x AM467913 (1-287)

QY 2 CysHsCysArGValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
Db 211 TGCTATTGCGAATACCAAGCTGTCATTGGAGGAAGTCNGCATGACCTGCACTAC 152
QY 22 LeuGlyGluArGlyTyrProIleCysCys 30
Db 151 CAGGGAAGACTCGGCGATCTGCTGC 125

RESULT 7
LOCUS BI021835 349 bp mRNA linear EST 14-JUN-2001
DEFINITION RCS-MT0259-170101-011-A09 MT0259 Homo sapiens cDNA, mRNA sequence.
ACCESSION BI021835
VERSION BI021835.1 GI:14428465
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 349) Dias Neto E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
JOURNAL
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RCS-MT0259-170101-011-A09&ts=2001-01-17&ta=1>
Seq primer: puc 18 forward
High quality sequence stop: 274.

FEATURES

source
1..349
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1b="MT0259"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles

Information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgl/hgl.html>)
Seq primer: M13 Reverse.

FEATURES

source

Location/Qualifiers
1..229
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="ATCC (inhost):122032"
/db_xref="taxon:9606"
/sex="mixed"
/tissue_type="bone marrow"
/dev_stage="adult"
/clone_lib="Bone marrow"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT 66 a 54 c 49 g 60 t
ORIGIN

Alignment Scores:

Pred. No.: 0.575 Length: 229
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 9 Gaps: 0

US-10-045-180A-6 (1-31) x AA321481 (1-229)

QY 2 CysHsCysArgValLeuTyrcysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
|||:::|||||:::|||||
Db 22 TGCATTGCGAGATACCGACGTCGATTGAGAGAACGTCGCTATGAACTGCACTTAC 81
|||:::|||||
QY 22 LeuGlyGluArgTyrcProIleCysCys 30
|||:::|||||
Db 82 CAGGGAAGACTCTGGGCACTTCTGCTGC 108
|||:::|||||

RESULT 12
AM075767 233 bp mRNA linear EST 13-OCT-1999
LOCUS AM075767/c
DEFINITION xasb06.x1 NCI CGAP CML1 Homo sapiens cDNA clone IMAGE:2573555 3'
similar to gb:M26602 NEUTROPHIL DEFENSINS 1, 2 AND 3 PRECURSOR (HUMAN); mRNA sequence.
ACCESSION AM075767
VERSION AM075767.1 GI:6030765
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 233)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Elisabeth Paietta, Jonathan D. Licht, M.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christina Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: www.bio.lnlnl.gov/bdrr/image/image.html
Seq primer: -40UP from Gibco
High quality sequence stop: 226.
Location/Qualifiers

FEATURES

source

1..233
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="IMAGE:2573555"
/issue_type="myeloid cells, 18 pooled CML cases, BCR/ABL"
Location/Qualifiers

rearrangement positive, includes both chronic phase and myeloid blast crisis"
/lab_host="DH10B"
/clone_lib="NCI CGAP CML1"
/note="Organ: whole blood; Vector: pCMV-Sport6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: 0190 dt. Library constructed by Life Technologies."
BASE COUNT 62 a 49 c 55 g 67 t
ORIGIN

Alignment Scores:

Pred. No.: 0.588 Length: 233
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 9 Gaps: 0

US-10-045-180A-6 (1-31) x AM075767 (1-233)

QY 2 CysHsCysArgValLeuTyrcysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
|||:::|||||:::|||||
Db 207 TGCATTGCGAGATACCGACGTCGATTGAGAGAACGTCGCTATGAACTGCACTTAC 148
|||:::|||||
QY 22 LeuGlyGluArgTyrcProIleCysCys 30
|||:::|||||
Db 147 CAGGGAAGACTCTGGGCACTTCTGCTGC 121
|||:::|||||

RESULT 13
BF893374/c 234 bp mRNA linear EST 18-JAN-2001
LOCUS BF893374/c
DEFINITION QV3-MT0129-111100-427-a02 MT0129 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF893374
VERSION BF893374.1 GI:12284833
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 234)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H., Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.J.
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
CONTACT: Simpson A.J.J.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL (<http://www.ludwig.org.br/scripts/gethtml2.pl?pl=QV3&t2=QV3-MT0129-111100-427-a02&t3=2000-11-11&t4=1>)
Seq primer: puc 18 forward
High quality sequence stop: 234.
Location/Qualifiers

FEATURES

source

1..234
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="MT0129"
/note="Organ: marrow; Vector: puc18; Site_1: SmaI; Site_2:
Location/Qualifiers

Smal: A mini-library was made by cloning products derived from ORSTS PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

BASE COUNT 69 a 55 c 57 g 65 t
ORIGIN

Alignment Scores:

Pred. No.:	0.633	Length:	246
Score:	80.00	Matches:	13
Percent Similarity:	55.17%	Conservative:	3
Best Local Similarity:	44.83%	Mismatches:	13
Query Match:	41.88%	Indels:	0
DB:	10	Gaps:	0

US-10-045-180A-6 (1-31) X BF903426 (1-246)

QY	2	CySHiCyArGValLeuTYrCysIIlePheGlyGluHisLeuGlyGlyThrCysPheIle	21
DB	66	TGCTATTGCAGATATCCAGCGTGCATTGCAGGAACTCGCTATGAACTGCATCTAC	125
QY	22	LeuGlyGluArgTYrProIleCysCys	30
DB	126	CAGGGAAGACTCTGGGCATTCTGCTGC	152

Search completed: December 17, 2003, 18:25:48
Job time : 1085.85 secs

PN FR2767832-A1.
XX 05-MAR-1999.
XX 29-AUG-1997; 97FR-0010823.
XX 29-AUG-1997; 97FR-0010823.
XX 29-AUG-1997; 97FR-0010823.
XX (GENST) GENSET SA.
XX Bougueleret L, Chumakov I;
XX WPI, 1999-183266/16.
XX P-PSDB; AAY01604.
XX Human defensin polypeptide Def-X - useful as antimicrobial agent,
XX anticancer agent, pesticide, etc.
XX Disclosure, Fig 3; 56pp; French.
XX The present sequence encodes human defensin (Def-X). The Def-X
XX polypeptide can be used as an antimicrobial, antiparasitic agent or
XX a pesticide. The Def-X polypeptide can be used as a cytostatic
XX (especially anticancer) agent, and as an agent for modulating processes
XX of inflammation, tissue repair and endocrine (especially corticostatic)
XX regulation. The polypeptide can be used in a composition for external
XX topical use, especially in a cosmetic composition. Compositions
XX containing Def-X can be used for prevention and treatment of microbial
XX and parasitic infections, especially where the microbial or parasitic
XX infections are Gram-positive or -negative bacterial infections or
XX mycobacterial, fungal or spirochaete infections, or where the viral
XX infections are associated with enveloped viruses, especially HSV and HIV.
XX The compositions can be used for prevention and/or treatment of cancers,
XX especially melanomas, or liver cancer, prostate cancer, non-small-cell
XX lung cancer or colorectal carcinoma, and for enhancing immunity,
XX especially in the case of AIDS, or preventing immune deficiency,
XX especially in the treatment of psoriasis, as well as for modulating
XX inflammatory processes, especially in the case of chronic inflammatory
XX disorders.
XX Sequence 453 BP; 107 A; 121 C; 97 G; 128 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 2.99e-17 Length: 453
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-045-180A-6 (1-31) x AAX26697 (1-453)
QY 1 IIECYHISCYAARGYVALLLEUTYRCYSIIIEPHGGLYGLNHSLENGLYGLYTHRCYSY 20
DB 241 ATCTGCCATTGCGAGACTACTGCTATTTTGGAGAACATCTGGTGGACCTGCTTC 300
QY 21 IIELENGLYGLUARGTYRPROILECYSYSTYR 31
DB 301 ATCTTGGTGAACGCTACCAATCTGCTGCTAC 333
RESULT 2
AAX26698
ID AAX26698 standard; cDNA; 453 BP.
XX
XX AAX26698;
XX
XX 18-UTN-1999 (first entry)
XX
XX cDNA sequence encoding human defensin (Def-X) protein.
XX Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
XX cytostatic; anticancer; inflammation; tissue repair;
XX endocrine regulation; corticostatic regulation; cancer; melanoma;
XX

KW AIDS; immune deficiency; psoriasis; ss.
XX
XX OS Homo sapiens.
XX
XX PN FR2767832-A1.
XX 05-MAR-1999.
XX 29-AUG-1997; 97FR-0010823.
XX 29-AUG-1997; 97FR-0010823.
XX 29-AUG-1997; 97FR-0010823.
XX (GENST) GENSET SA.
XX Bougueleret L, Chumakov I;
XX WPI, 1999-183266/16.
XX
XX Human defensin polypeptide Def-X - useful as antimicrobial agent,
XX anticancer agent, pesticide, etc.
XX Claim 5; Fig 4; 56pp; French.
XX The present sequence encodes human defensin (Def-X). The Def-X
XX polypeptide can be used as an antimicrobial, antiparasitic agent or
XX a pesticide. The Def-X polypeptide can be used as a cytostatic
XX (especially anticancer) agent, and as an agent for modulating processes
XX of inflammation, tissue repair and endocrine (especially corticostatic)
XX regulation. The polypeptide can be used in a composition for external
XX topical use, especially in a cosmetic composition. Compositions
XX containing Def-X can be used for prevention and treatment of microbial
XX and parasitic infections, especially where the microbial or parasitic
XX infections are Gram-positive or -negative bacterial infections or
XX mycobacterial, fungal or spirochaete infections, or where the viral
XX infections are associated with enveloped viruses, especially HSV and HIV.
XX The compositions can be used for prevention and/or treatment of cancers,
XX especially melanomas, or liver cancer, prostate cancer, non-small-cell
XX lung cancer or colorectal carcinoma, and for enhancing immunity,
XX especially in the case of AIDS, or preventing immune deficiency,
XX especially in the treatment of psoriasis, as well as for modulating
XX inflammatory processes, especially in the case of chronic inflammatory
XX disorders.
XX Sequence 453 BP; 107 A; 121 C; 97 G; 128 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 2.99e-17 Length: 453
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-10-045-180A-6 (1-31) x AAX26698 (1-453)
QY 1 IIECYHISCYAARGYVALLLEUTYRCYSIIIEPHGGLYGLNHSLENGLYGLYTHRCYSY 20
DB 241 ATCTGCCATTGCGAGACTACTGCTATTTTGGAGAACATCTGGTGGACCTGCTTC 300
QY 21 IIELENGLYGLUARGTYRPROILECYSYSTYR 31
DB 301 ATCTTGGTGAACGCTACCAATCTGCTGCTAC 333
RESULT 3
AAX26696
ID AAX26696 standard; DNA; 4415 BP.
XX
XX AAX26696;
XX
XX 18-UTN-1999 (first entry)
XX
XX Nucleotide sequence of human defensin (Def-X).
XX

DB: 64 CAGGTCGTCTGTGGCGTTCTGCTGC 90

RESULT 5
ID AAT17899 standard; DNA; 108 BP.
AAT17899;
AC AAT17899;
DT 24-SEP-1996 (first entry)
XX
DE DNA encoding amphiphilic peptide human NP-1.
XX
KW Human NP-1; amphiphilic peptide; recombinant production;
KW protease deficient; microbial host cell; expression vector;
KW *Baccharichia coli*; K-12 cell; vector; cleavable fusion protein;
KW carbohydrate binding protein; anti-parasitic; anti-fungal;
KW anti-tumour; anti-cancer; anti-viral; anti-microbial; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 6..101
FT /*tag= a
XX
PN WO9604373-A2.
XX
PD 15-FEB-1996.
XX
PF 26-JUL-1995; 95WO-US10219.
XX
PR 29-JUL-1994; 94US-0282030.
XX
PA (MAGA-) MAGANIN PHARM INC.
XX
PI Anderson GM, Kari P, Pierce JC, Williams JT;
XX
DR WPI; 1996-129390/13.
DR P-PSDB; AAR92824.
XX
XX
PT Recombinant production of amphiphilic peptide in protease deficient
PT microbial host, pref. *E. coli* K-12 - useful in prodn. of
PT antimicrobial, antiviral and anticancer peptide(s)
XX
PS Disclosure; Page 24; 103pp; English.
XX
XX The present sequence, which encodes human NP-1 (an amphiphilic
CC peptide) is used in 2 novel methods for the recombinant prodn. of
CC human NP-1. The 1st method comprises transforming a protease
CC deficient (PD) microbial host cell with an expression vector contg.
CC the DNA, under the control of a regulatory sequence operable in the
CC host, and expressing the peptide in the transformed host. The 2nd
CC method comprises transforming an *E. coli* PD K-12 cell with a vector
CC that expresses a cleavable fusion protein, comprising at least part
CC of a carbohydrate binding protein (CBP) and the peptide, expressing
CC the fusion protein in the cell and cleaving the protein to obtain
CC the peptide substantially free of CBP residues. These methods for
CC producing and processing human NP-1 allow high levels of the
CC peptide to accumulate in certain PD microbial host cells, despite
CC the peptide's anti-microbial potency, and efficient recovery of the
CC full length peptide. The peptide produced, unlike most natural
CC analogous peptides, exhibits a broader range of activity and/or
CC greater potency compared to a related natural peptide, e.g. the purf
CC peptide may be used as an anti-parasitic, anti-fungal, anti-tumour,
CC anti-cancer or an anti-viral agent.
XX
SQ Sequence 108 BP; 20 A, 25 C, 28 G, 35 T, 0 other;

Alignment Scores:
Pred. No.: 0.00977 Length: 108
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0

DB: 17 Gaps: 0

US-10-045-180A-6 (1-31) x AAT17899 (1-108)

QY 2 CysHisGValGValLeuTrpCysIlePheGlyIuHsleuGlyGlyThrCysPheIle 21
DB 12 TGTACTGCGCGATTTCGCGATCGACGCGAGCGTCTGCTACTGATTATTC 71
QY 22 LeuGlyGluArgTrpProIleCysCys 30
DB 72 CAGGTCGTCTGTGGCGATTCTGTTGC 98

RESULT 6
ID AA229363 standard; DNA; 110 BP.
AA229363;
AC AA229363;
XX
DT 29-FEB-2000 (first entry)
XX
DE Antimicrobial peptide, human alpha defensin encoding DNA.
XX
KW purf gene; glutamine, pyrophosphoribosyl pyrophosphate amidotransferase;
KW purf derivative; fusion partner; antimicrobial peptide; alpha defensin;
KW mass production; cleavage site; hydroxylamine; CNBr; DNA construct;
KW neutralise; toxicity; pharmaceutical industry; food industry; de.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 11..103
FT /*tag= a
FT /product= "Human alpha defensin peptide"
FT /note= "Antimicrobial peptide used in DNA construct"
XX
PN WO964611-A1.
XX
PD 16-DEC-1999.
XX
PP 08-JUN-1999; 99WO-KR00282.
XX
PR 09-JUN-1998; 98KR-0022117.
PR 14-MAY-1999; 99KR-0017920.
XX
PA (SAMY-) SAMYANG GENEX CORP.
XX
PI Kim JH, Kang MH, Lee J, Park SH, Lee JW, Hong SS, Lee H;
XX
DR WPI; 2000-097542/08.
DR P-PSDB; AAY44323.
XX
XX
PT New DNA constructs useful for mass production of antimicrobial peptides
PT in microorganism hosts -
XX
PS Claim 1; Fig 1; 67pp; English.
XX
XX The present DNA sequence encodes an antimicrobial peptide, human alpha
CC defensin. It is used along with a derivative of purf gene sequence that
CC functions as a fusion partner. A DNA construct that comprises, this
CC antimicrobial peptide encoding sequence and the entire, partial or
CC derivative of purf gene, is used for mass production of the antimicrobial
CC peptide in microorganisms without killing the host cells. Use of the purf
CC gene derivative sequence, neutralises the toxicity of the antimicrobial
CC peptides against the host microorganism. The antimicrobial peptides are
CC useful commercially in the pharmaceutical and food industries.
XX
SQ Sequence 110 BP; 14 A, 31 C, 37 G, 28 T, 0 other;

Alignment Scores:
Pred. No.: 0.00998 Length: 110
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13

Query Match: 41.88% Indels: 0
DB: 21 Gaps: 0
US-10-045-180A-6 (1-31) x AA229363 (1-110)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
DB 14 TGTACTATGCCGATATCCCGCGTGCATCCGGGTGAGCGTTCATTACCGTACCTGCATCTAC 73
QY 22 LeuGlyGluArgTyrProIleCysCys 30
DB 74 CAGGTCGTCTGTGGCGCTTCGCTGC 100

RESULT 7
ABL60444
ID ABL60444 standard; DNA; 110 BP.
AC ABL60444;
XX
XX 28-MAR-2003 (first entry)
DT
XX
XX DNA fragment of the invention HNP-1.
DE
XX
XX Gene expression; peptide antibiotic; purf gene; ds.
KM
XX
XX Unidentified.
OS
XX
XX Key Location/Qualifiers
FH 11.103
FT CDS /*tag= a
FT /partial
FT /note= "no start codon present"
XX
XX KR2001098973-A.
XX
XX PD 08-NOV-2001.
XX
XX PF 08-JUN-2001; 2001KR-0031889.
XX
XX PR 08-JUN-2001; 2001KR-0031889.
XX
XX PA (SAMV-) SAMYANG GENEX CORP.
XX
XX PI Hong SS, Kang WH, Kim JH, Lee HS, Lee JH, Lee JW, Park SH;
XX WPI; 2002-301977/34.
XX DR P-PSDB; ABB81939.
XX
XX PT Gene expression system useful for mass-production of peptide
PT antibiotics and vectors derived from microorganisms -
XX
XX PS Disclosure; Page 16; 56pp; Korean.
XX
XX The invention relates to a genetic component which mass-produces peptide
XX antibiotics effectively from microorganisms. The genetic component
XX consists of a first gene sequence which codes for the whole or partial
XX purf gene or its derivative, and a second gene sequence which codes
XX peptide antibiotics. The mass-production method of peptide antibiotics
XX comprises the steps of: constructing an expression vector including the
XX genetic component, transforming a bacterial host cell with the vector,
XX culturing the transformed cell to express the genetic component, and
XX recovering the peptide antibiotics. The expression vector is selected
XX from the group consisting of pGNX2, pGNX3, pGNX4 and pGNX5, and it has
XX a high copy number of origin, strong transcription promoter and
XX structural gene. The sequences given in records ABL60400-ABL60464
XX represent DNA sequences of the invention.
SQ Sequence 110 BP; 14 A; 31 C; 37 G; 28 T; 0 other;

Alignment Scores:
Pred. No.: 0.00998 Length: 110
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3

Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 24 Gaps: 0
US-10-045-180A-6 (1-31) x ABL60444 (1-110)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
DB 14 TGTACTATGCCGATATCCCGCGTGCATCCGGGTGAGCGTTCATTACCGTACCTGCATCTAC 73
QY 22 LeuGlyGluArgTyrProIleCysCys 30
DB 74 CAGGTCGTCTGTGGCGCTTCGCTGC 100

RESULT 8
AA188805
ID AA188805 standard; cDNA; 159 BP.
XX
XX AA188805;
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human polynucleotide SEQ ID NO 8865.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KM vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KM tissue growth factor; immunomodulatory; cancer; leukemia;
KM nervous system disorders; arthritis; inflammation; BS.
XX
XX OS Homo sapiens.
XX
XX PN WO200164835-A2.
XX
XX PD 07-SEP-2001.
XX
XX PF 26-FEB-2001; 2001WO-US04927.
XX
XX PR 28-FEB-2000; 2000US-0515126.
XX
XX PR 18-MAY-2000; 2000US-0577409.
XX
XX PA (HYSE-) HYSEQ INC.
XX
XX PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX DR P-PSDB; AAO08874.
XX
XX PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukemia, inflammation and immune
PT disorders -
XX
XX PS Claim 1; SEQ ID NO 8865; 1399bp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AA179941-AA193841) and
XX the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.

Alignment Scores:
Pred. No.: 0.0156 Length: 159
Score: 80.00 Matches: 13

Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: 22 Gaps: 0

US-10-045-180a-6 (1-31) x AAF188805 (1-159)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
 DB 15 TGCATTATTCAGAAATACACGAGCGTGCATTGCAGAGAAACGTCGTATGGAACTGCATCTAC 74

QY 22 LeuGlyGluArgTyrProIleCysCys 30
 DB 75 CAGGGAAGACTCTGGGCAATTCCTGCTGC 101

RESULT 9
 AAF20911
 ID AAF20911 standard; DNA; 464 BP.
 AC AAF20911;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 DE Human defensin 3 polynucleotide fragment #2478.
 XX
 KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antiaesthetic; analgesic; hypotensive; cytostatic;
 KW respiratory obstruction; pulmonary; analgesic; hypotensive; cytostatic;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 KW
 OS Homo sapiens.
 XX
 PN WO200062736-A2.
 XX
 PD 26-OCT-2000.
 XX
 PF 24-MAR-2000; 2000WO-US08020.
 XX
 PR 06-APR-1999; 99US-0127958.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 PA (NYCE/) NYCE J W.
 XX
 PI Nyce JW;
 XX
 DR WPI; 2000-679539/66.
 XX
 PT Low adenosine (A) content antisense oligonucleotides which do not
 PT trigger adenosine receptors during metabolism, useful e.g. for treating
 PT cancers and respiratory obstructions -
 XX
 PS Disclosure; Page 218; 1592pp; English.
 XX
 CC The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antiaesthetic, hypotensive and cytostatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine,
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system

CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and
 CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impaired respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.
 CC
 SQ Sequence 464 BP; 110 A; 128 C; 119 G; 107 T; 0 other;

Alignment Scores:
 Pred. No.: 0.0564 Length: 464
 Score: 80.00 Matches: 13
 Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: 21 Gaps: 0

US-10-045-180a-6 (1-31) x AAF20911 (1-464)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
 DB 285 TGCATTATTCAGAAATACACGAGCGTGCATTGCAGAGAAACGTCGTATGGAACTGCATCTAC 344

QY 22 LeuGlyGluArgTyrProIleCysCys 30
 DB 345 CAGGGAAGACTCTGGGCAATTCCTGCTGC 371

RESULT 10
 AAA34789
 ID AAA34789 standard; DNA; 464 BP.
 AC AAA34789;
 XX
 DT 28-JUL-2000 (first entry)
 XX
 DE Human adenosine receptor related polynucleotide SEQ ID NO:2478.
 XX
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiaesthetic; cytostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
 KW
 OS Homo sapiens.
 XX
 PN WO200009525-A2.
 XX
 PD 24-FEB-2000.
 XX
 PF 03-AUG-1999; 99WO-US17712.
 XX
 PR 03-AUG-1998; 98US-0095212.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 PA
 PI Nyce JW;
 XX
 DR WPI; 2000-205971/18.
 XX
 PT New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,

PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers -
 XX
 PS Disclosure; Page 612; 1343pp; English.
 CC
 CC The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytosstatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airflow, including lung disease and diseases whose secondary
 CC effects affect the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impeded respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,
 CC carcinomas, and cancers which may metastasize to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing the
 CC bronchoconstriction and inflammation. AAA32313 to AAA5312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA3223 to AAA3392) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.
 XX
 SQ Sequence 464 BP; 110 A; 128 C; 119 G; 107 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.0564 Length: 464
 Score: 80.00 Matches: 13
 Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: Gaps: 0
 US-10-045-180a-6 (1-31) x AAA34789 (1-464)
 QY 2 CysHIScysArgValLeuTyrCys11PheGlyGluHisLeuGlyGlyThrCysPhe1le 21
 DB 285 TGTATTGCAAGAAATACACGCGTCGATTGCAGAGAACGTCGTATGGAACCTGCATCTAC 344
 QY 22 LeuGlyGluArgTyrPro1leCysCys 30
 DB 345 CAGGGAAGACTCTGGGCAATTCTGCTGC 371
 RESULT 11
 AAC04468
 ID AAC04468 standard; cDNA, 478 BP.
 XX
 AC AAC04468;
 XX
 DT 06-OCT-2000 (first entry)
 XX
 DE Human secreted protein 5' EST. SEQ ID NO: 8543.
 XX
 KM Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KM gene therapy; chromosome mapping; ss.
 XX
 OS Homo sapiens.
 XX
 PN EPI033401-A2.
 XX
 PD 06-SEP-2000.
 XX
 PF 21-FEB-2000; 2000EP-0200610.
 XX
 PR 26-FEB-1999; 99US-0122487.

XX
 PA (GEST) GENSER.
 XX
 PI Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 DR WPI; 2000-500381/45.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 PS Claim 1; SEQ ID 8543; 71pp + CD-ROM; English.
 XX
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
 CC identified within the present sequence. The 5' ESTs were prepared from
 CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
 CC sequences usually correspond mainly to the 3' untranslated region (UTR)
 CC of the mRNA because they are not well suited for isolating cDNA sequences
 CC derived from the 5' ends of mRNAs and even in those cases where longer
 CC cDNA sequences have been obtained, the full 5' UTR is rarely included.
 CC 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be
 CC used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used
 CC in diagnostic, forensic, gene therapy and chromosome mapping procedures.
 CC They are used to obtain upstream regulatory sequences and to design
 CC expression and secretion vectors.
 XX
 SQ Sequence 478 BP; 117 A; 129 C; 127 G; 101 T; 4 other;
 Alignment Scores:
 Pred. No.: 0.0585 Length: 478
 Score: 80.00 Matches: 13
 Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: Gaps: 0
 US-10-045-180a-6 (1-31) x AAC04468 (1-478)
 QY 2 CysHIScysArgValLeuTyrCys11PheGlyGluHisLeuGlyGlyThrCysPhe1le 21
 DB 346 TGTATTGCAAGAAATACACGCGTCGATTGCAGAGAACGTCGTATGGAACCTGCATCTAC 405
 QY 22 LeuGlyGluArgTyrPro1leCysCys 30
 DB 406 CAGGGAAGACTCTGGGCAATTCTGCTGC 432
 RESULT 12
 AAQ14908
 ID AAQ14908 standard; cDNA, 498 BP.
 XX
 AC AAQ14908;
 XX
 DT 25-FEB-1992 (first entry)
 XX
 DE Chronic myelogenous leukaemia-derived mrs cDNA.
 XX
 KM CML; Philadelphia chromosome; myeloid-related sequence;
 KM chromosome 8; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT CDS Location/Qualifiers
 FT 16..378
 FT /*tag= a
 FT /*note= "see comments"
 FT polyA_signal
 FT 479..484
 FT /*tag= b
 XX
 PN US5066792-A.
 XX
 PD 19-NOV-1991.

```
XX PF 02-DEC-1988; 88US-0279451.
XX PR 02-DEC-1988; 88US-0279451.
XX PR 28-SEP-1984; 84US-0655942.
XX PR 04-SEP-1987; 87US-0094099.
XX
XX (TEXA ) UNIT OF TEXAS.
XX
XX Saunders GF, Mars WM;
XX
XX MPI: 1991-361533/49.
XX DR P-PSDB; AARI5222.
XX
XX PT Gene probe for chronic myelogenous leukaemia - esp. useful for
XX distinguishing myeloblastic crisis from lymphoblastic crisis
XX
XX PS Disclosure; Fig 2; 13pp; English.
XX
XX CC The precise position of the CDS is unsure due to the 13 unidentified
XX nucleotides at the 5' end of the sequence. The most likely position
XX for initiation of translation occurs at nucleotide 94 and the open
XX reading frame following this point is claimed. The first 18 codons
XX of the claimed ORF (excluding the ATG) encode a sequence of amino
XX acids that resembles other leader sequences and their cleavage sites.
XX CC The mrs sequence was isolated as clone C-A3 (ATCC 39868) from a cDNA
XX library constructed from total RNA from a patient suffering from
XX chronic phase, Ph+-positive CML. The mrs is localised to human
XX chromosome 8, specifically to bands 8q21.1-23.
XX
XX SQ Sequence 498 BP; 118 A; 134 C; 122 G; 111 T; 13 other;

Alignment Scores:
Pred. No.: 0.0614 Length: 498
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 12 Gaps: 0

US-10-045-180A-6 (1-31) x AAQ14908 (1-498)
QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
DB 289 TGCATTCGACAAATACCAAGCGTCATTCGAGAGAACTGCTATGAACTGCATCTAC 348
QY 22 LeuGlyGluArgTyrProIleCysCys 30
DB 349 CAGGGAAGACTCTGGGCATTCTGCTGC 375

RESULT 13
ABQ80601
ID ABQ80601 standard; DNA; 498 BP.
XX
XX AC ABQ80601;
XX
XX DT 11-NOV-2002 (first entry)
XX
XX DE Human defensin 1 coding sequence.
XX
XX KW Antiallergic; allergic disease; allergy; defensin 1; atopic dermatitis;
XX human; gene; ds.
XX
XX OS Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT CDS 94..378
XX FT /*tag= a
XX FT /product= "Human defensin 1"
XX
XX PN WO200264832-A1.
XX
XX PD 22-AUG-2002.
```

```
XX PF 13-FEB-2002; 2002MO-JP01193.
XX PR 14-FEB-2001; 2001JP-0036362.
XX
XX PA (GENO-) GENOX RES INC.
XX PA (NIGR-) JAPAN GEN AGENCY NATION.
XX
XX Sugita Y, Hetshi M, Kagaya S, Gunji S, Satco H;
XX
XX MPI: 2002-674884/72.
XX DR P-PSDB; ABB98494.
XX
XX PT Examining allergic diseases by changes in expression levels of
XX defensin1 gene in peripheral blood monocytes as indication, also
XX applicable in screening compounds for treating of allergic diseases
XX e.g. atopic dermatitis
XX
XX PS Disclosure; Page 45-46; 57pp; Japanese.
XX
XX CC The present invention relates to a method for examining allergic
XX diseases. The method comprises: (a) determining the expression level of
XX defensin 1 gene in the biological sample from a patient; and (b)
XX comparing the expression level with that in the sample of a healthy
XX individual. The method is for examining allergic diseases particularly
XX atopic dermatitis and its diagnosis, which is also applicable in
XX screening candidate compounds for remedies. The present sequence is the
XX coding sequence for human defensin 1.
XX
XX SQ Sequence 498 BP; 118 A; 134 C; 122 G; 111 T; 13 other;

Alignment Scores:
Pred. No.: 0.0614 Length: 498
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 24 Gaps: 0

US-10-045-180A-6 (1-31) x ABQ80601 (1-498)
QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
DB 289 TGCATTCGACAAATACCAAGCGTCATTCGAGAGAACTGCTATGAACTGCATCTAC 348
QY 22 LeuGlyGluArgTyrProIleCysCys 30
DB 349 CAGGGAAGACTCTGGGCATTCTGCTGC 375

RESULT 14
AAC00098
ID AAC00098 standard; cDNA; 509 BP.
XX
XX AC AAC00098;
XX
XX DT 06-OCT-2000 (first entry)
XX
XX DE Human secreted protein 5' EST, SEQ ID NO: 96.
XX
XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping; ss.
XX
XX OS Homo sapiens.
XX
XX PN EP1033401-A2.
XX
XX PD 06-SEP-2000.
XX
XX PF 21-FEB-2000; 2000EP-0200610.
XX
XX PR 26-FEB-1999; 99US-0122487.
XX
XX PA (GIST ) GENSET.
```

XX Dumas Milne Edwards J, Duclert A, Giordano J;
 XX
 PI MPI: 2000-500381/45.
 DR P-PSDB; AAC00092.
 XX
 PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 1; SEQ ID 96; 71pp + CD-ROM; English.
 XX
 CC The present sequence is one of a large number of 5' ESTs derived from
 CC mRNAs encoding secreted proteins. An ORF has been identified within the
 CC sequence. The 5' ESTs were prepared from total human RNAs or poly(A)⁺ RNAs
 CC derived from 30 different tissues. EST sequences usually correspond
 CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
 CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
 CC well suited for isolating cDNA sequences derived from the 5' ends of
 CC mRNAs and even in those cases where longer cDNA sequences have been
 CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
 CC cDNAs with intact 5' ends and can therefore be used to obtain full length
 CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
 CC gene therapy and chromosome mapping procedures. They are used to obtain
 CC upstream regulatory sequences and to design expression and secretion
 CC vectors.
 XX
 SQ Sequence 509 BP; 118 A; 145 C; 132 G; 109 T; 5 other;
 XX
 Alignment Scores:
 Pred. No.: 0.063 Length: 509
 Score: 80.00 Matches: 13
 Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: 21 Gaps: 0
 XX
 US-10-045-180a-6 (1-31) x AAC00098 (1-509)
 QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
 Db 377 TGTATTGCAAGAAATACACAGCGTGCAATTGCAGAGAACGTCGCTATGAACTGCACTAC 436
 QY 22 LeuGlyGluArgTyrProIleCysGys 30
 Db 437 CAGGGAAGACTCTGGGCAATTCGTCTGC 463
 XX
 RESULT 15
 ABNS9722
 ID ABNS9722 standard; cDNA; 514 BP.
 XX
 AC ABNS9722;
 XX
 DT 28-JUN-2002 (first entry)
 XX
 DE Novel human coding sequence SEQ ID NO: 133.
 XX
 KM Human; anti-nausea; vulnerary; anti-inflammatory; immunomodulator;
 KM antifertility; cerebroprotective; cytosolic; rheumatic; gene therapy;
 KM neuroprotective; antiparkinsonian; protein therapy; EST;
 KM expressed sequence tag; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200222660-A2.
 XX
 PD 21-MAR-2002.
 XX
 PF 10-SEP-2001; 2001WO-US26015.
 XX
 PR 11-SEP-2000; 2000US-0659671.
 XX
 PA (HYSR-) HYSRQ INC.

XX Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Wehrman T, Drmanac RR;
 XX
 DR MPI: 2002-292408/33.
 DR P-PSDB; ABB97309.
 XX
 PT An isolated polynucleotide for treating diseases associated with its
 PT encoded polypeptide such as cancer and multiple sclerosis -
 XX
 PS Claim 1; SEQ ID NO 133; 509pp; English.
 XX
 CC The present invention provides the protein and coding sequences of 444
 CC novel human proteins. These were isolated from expressed sequences tags
 CC (ESTs). They can be used to stimulate cell growth to regulate
 CC haematopoiesis e.g. to treat aplastic anaemia, to help tissue regrowth
 CC e.g. in burn treatment, to regulate the immune system e.g. to treat
 CC multiple sclerosis, to regulate activity or inhibit e.g. to treat
 CC infertility, to regulate haemostasis or thrombolysis e.g. to treat
 CC stroke and cancer, to screen for drugs, to treat inflammatory conditions
 CC e.g. rheumatoid arthritis, and to treat nervous system disorders e.g.
 CC Parkinson's disease. The present sequence is a coding sequence of the
 CC invention.
 XX
 SQ Sequence 514 BP; 138 A; 142 C; 118 G; 116 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.0638 Length: 514
 Score: 80.00 Matches: 13
 Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: 24 Gaps: 0
 XX
 US-10-045-180a-6 (1-31) x ABNS9722 (1-514)
 QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
 Db 292 TGTATTGCAAGAAATACACAGCGTGCAATTGCAGAGAACGTCGCTATGAACTGCACTAC 351
 QY 22 LeuGlyGluArgTyrProIleCysGys 30
 Db 352 CAGGGAAGACTCTGGGCAATTCGTCTGC 378
 XX
 Search completed: December 17, 2003, 15:06:05
 Job time : 118.25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: December 17, 2003, 15:06:18 ; Search time 107.181 Seconds
(without alignments)
963.776 Million cell updates/sec

Title: US-10-045-180A-6
Perfect score: 191
Sequence: 1 ICHCRVLYCIGFHEHLGTCFFILGERYPICCY 31

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2211978 segs, 1666101734 residues
Total number of hits satisfying chosen parameters: 4423956

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frames.p2n.model -DRV=xld
-Q=/cgn2_1/USPTO.spool_p/US10045180/runat_17122003_145026_12635/app_query.fasta_1.860
-DB=Published Applications_NA -QFMT=fastcap -SUFFI=rmdb -MINMATCH=0.1
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62
-TRANS-human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pic -NOR=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10045180 @CGN 1.1 174 @runat_17122003_145026_12635
-NCPU=6 -ICPU=3 -NO MMAP -LARGESUBSTRY -NEG_SCORES=0 -WAIT -DSRBLCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/prodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/prodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq2:*
14: /cgn2_6/prodata/2/pubpna/US10A_PUBCOMB.seq:*
15: /cgn2_6/prodata/2/pubpna/US10B_PUBCOMB.seq:*
16: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*
17: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*
18: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Result Query Match Length DB ID Description
SIMMARIES

1	191	100.0	453	14	US-10-013-770-2	Sequence 2, Appli
2	191	100.0	453	14	US-10-045-180A-2	Sequence 2, Appli
3	191	100.0	4415	14	US-10-013-770-1	Sequence 1, Appli
4	191	100.0	4415	14	US-10-045-180A-1	Sequence 1, Appli
5	80	41.9	337	11	US-09-918-995-18198	Sequence 18198, A
6	80	41.9	337	11	US-09-918-995-19070	Sequence 19070, A
7	80	41.9	353	11	US-09-918-995-18323	Sequence 18323, A
8	80	41.9	555	11	US-10-252-157-223	Sequence 223, App
9	80	41.9	726	10	US-09-816-828-16	Sequence 16, Appl
10	80	41.9	726	10	US-09-816-828-16	Sequence 16, Appl
11	80	41.9	726	10	US-09-816-828-16	Sequence 78, Appl
12	80	41.9	1348	15	US-10-125-237-18	Sequence 78, Appl
13	80	41.9	1348	15	US-10-125-237-18	Sequence 1047, Ap
14	74	38.7	412	10	US-09-902-941-1047	Sequence 1047, Ap
15	74	38.7	412	10	US-09-849-626-1047	Sequence 1047, Ap
16	74	38.7	412	10	US-10-113-872-1047	Sequence 1047, Ap
17	74	38.7	412	10	US-10-017-754-1047	Sequence 6, Appli
18	74	38.7	967	13	US-10-162-043-6	Sequence 18788, A
19	69	36.1	343	11	US-09-918-995-18788	Sequence 18788, A
20	67	35.1	258	11	US-10-029-386-18673	Sequence 4917, Ap
21	67	35.1	588	13	US-10-029-386-18673	Sequence 18673, Ap
22	66.5	34.8	640	13	US-10-027-632-106465	Sequence 106465,
23	66.5	34.8	640	13	US-10-027-632-106465	Sequence 106466,
24	66.5	34.8	640	13	US-10-027-632-106465	Sequence 106465,
25	66.5	34.8	640	14	US-10-027-632-106465	Sequence 106466,
26	66.5	34.8	640	14	US-10-027-632-106465	Sequence 106467,
27	66.5	34.8	640	14	US-10-027-632-106467	Sequence 22769, A
28	66.5	34.8	738	13	US-10-027-632-22769	Sequence 22770, A
29	66.5	34.8	738	13	US-10-027-632-22770	Sequence 22771, A
30	66.5	34.8	738	14	US-10-027-632-22771	Sequence 22769, A
31	66.5	34.8	738	14	US-10-027-632-22770	Sequence 22770, A
32	66.5	34.8	738	14	US-10-027-632-22771	Sequence 22771, A
33	66.5	34.8	738	14	US-10-027-632-22771	Sequence 165609,
34	65	34.0	542	13	US-10-045-180A-8	Sequence 165609,
35	65	34.0	784	14	US-10-027-632-165609	Sequence 25141, A
36	65	34.0	787	13	US-10-027-632-25141	Sequence 25141, A
37	65	34.0	787	13	US-10-027-632-25142	Sequence 25141, A
38	65	34.0	787	13	US-10-027-632-25143	Sequence 25143, A
39	65	34.0	787	14	US-10-027-632-25143	Sequence 25143, A
40	65	34.0	787	14	US-10-027-632-25143	Sequence 25143, A
41	65	34.0	787	14	US-10-027-632-25143	Sequence 473, App
42	65	34.0	860	14	US-10-044-099-473	Sequence 7, Appli
43	65	34.0	860	14	US-10-044-099-473	Sequence 35, Appli
44	65	34.0	4295	11	US-10-045-180A-7	
45	63	33.0	15456	11	US-09-900-112-35	

ALIGNMENTS

RESULT 1
US-10-013-770-2
; Sequence 2, Application US/10013770
; Publication No. US20020115151A1
GENERAL INFORMATION:

APPLICANT: GENSET SA

TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
THERAPEUTIC APPLICATIONS

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: Knodde, Martens, Olson & Bear

STREET: 550 West C Street
CITY: San Diego
STATE: California

COUNTRY: USA
ZIP: 92101

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/013, 770

FILING DATE: 10-Dec-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/486,580
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.064C1
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 453 BASE PAIRS
TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: CDNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-10-013-770-2

Alignment Scores:
Pred. No.: 4.36e-19 Length: 453
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-045-180A-6 (1-31) x US-10-013-770-2 (1-453)

Qy 1 lleCysHsCysArGValleuTyrcysillepHeGlyGluHileuGlyGlyThrcysPhe 20
Db 241 ATTCGCCATTGCAGAGTCTATCTGCGATTTTGGAGAACATCTTGGTGGACCTGCTTC 300

Qy 21 lleleuGlyGluArgTyrcProlleCysCysTyrc 31
Db 301 ATCCTTGTAACGCTACCACTGCTGCTAC 333

RESULT 2
US-10-045-180A-2
Sequence 2, Application US/10045180A
Publication No. US20020182703A1
GENERAL INFORMATION:
APPLICANT: Bougueleret, Lydie
TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Composit
FILE REFERENCE: GEN-100D1
CURRENT APPLICATION NUMBER: US/10/045,180A
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: US 09/486,580
PRIOR FILING DATE: 2000-02-25
PRIOR APPLICATION NUMBER: PCT/FR98/01864
PRIOR FILING DATE: 1998-08-28
PRIOR APPLICATION NUMBER: FR 97/10823
PRIOR FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 453
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (52)..(336)
OTHER INFORMATION: Def-X coding sequence
US-10-045-180A-2

Alignment Scores:
Pred. No.: 4.36e-19 Length: 453
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-045-180A-6 (1-31) x US-10-045-180A-2 (1-453)

Qy 1 lleCysHsCysArGValleuTyrcysillepHeGlyGluHileuGlyGlyThrcysPhe 20
Db 241 ATTCGCCATTGCAGAGTCTATCTGCGATTTTGGAGAACATCTTGGTGGACCTGCTTC 300

Qy 21 lleleuGlyGluArgTyrcProlleCysCysTyrc 31
Db 301 ATCCTTGTAACGCTACCACTGCTGCTAC 333

RESULT 3
US-10-013-770-1
Sequence 1, Application US/10013770
Publication No. US20020115151A1
GENERAL INFORMATION:
APPLICANT: GENSET SA
TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
THERAPEUTIC APPLICATIONS

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESSES:
ADDRESSER: Knobbe, Martens, Olson & Bear
STREET: 550 West C Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/013,770
FILING DATE: 10-Dec-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/486,580
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.064C1
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4415 BASE PAIRS
TYPE: NUCLEOTIDE
STRANDEDNESS: DOUBLE
TOPOLOGY: LINEAR
MOLECULE TYPE: DNA
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: Exon 1
LOCATION: 1836..1874
FEATURE:
NAME/KEY: Exon 2
LOCATION: 3394..3577
FEATURE:
NAME/KEY: Exon 3
LOCATION: 4161..4380
FEATURE:
NAME/KEY: start CDS
LOCATION: 3406..3408
FEATURE:
NAME/KEY: stop CDS
LOCATION: 4276..4278
FEATURE:
NAME/KEY: polyadenylation site
LOCATION: 4374..4379
SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-10-013-770-1

Alignment Scores:

Pred. No.:	7,29e-18	Length:	4415
Score:	191.00	Matches:	31
Best Local Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	14	Gaps:	0

US-10-045-180A-6 (1-31) x US-10-013-770-1 (1-4415)

QY 111eCyBH1sCyatrgyValLeuTyrcys11lePheGly1uH1leuGlyGlyThrcysPhe 20

DB 4181 ATCTGCATTGACAGACTACTACTACTATTTTGGAGAACATCTGTGGACCTGCTTC 4240

QY 21 11e1euglyGlyuArgTyrrPro11eCyGcTyrr 31

DB 4241 ATCTTGTGTGAACGCTACCCACATCTGCTCTAC 4273

RESULT 4

US-10-045-180A-1

Sequence 1, Application US/10045180A

Publication No. US20020182703A1

GENERAL INFORMATION:

APPLICANT: Bougueleret, Lydie

APPLICANT: Chumakov, Ilya

TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Compositions Containing Them and Applications to Diagnosis and to Therapeutic

FILE REFERENCE: GEN-100D1

CURRENT APPLICATION NUMBER: US/10/045,180A

PRIOR FILING DATE: 2001-10-18

PRIOR APPLICATION NUMBER: US 09/486,580

PRIOR FILING DATE: 2000-02-25

PRIOR APPLICATION NUMBER: PCT/FR98/01864

PRIOR FILING DATE: 1998-08-28

PRIOR APPLICATION NUMBER: FR 97/10823

PRIOR FILING DATE: 1997-08-29

NUMBER OF SEQ ID NOS: 12

SOFTWARE: PatentIn version 3.1

SEQ ID NO 1

LENGTH: 4415

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1)..(4415)

OTHER INFORMATION: Def-X genomic sequence

FEATURE:

NAME/KEY: misc_feature

LOCATION: (85)..(85)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (143)..(143)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (670)..(670)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (970)..(970)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1111)..(1111)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1150)..(1150)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: CAAT signal

LOCATION: (1711)..(1714)

OTHER INFORMATION:

FEATURE:

NAME/KEY: TATA signal

LOCATION: (1758)..(1767)

OTHER INFORMATION:

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1780)..(1780)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1836)..(1874)

OTHER INFORMATION: Exon 1

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1875)..(1880)

OTHER INFORMATION: splice donor site

FEATURE:

NAME/KEY: misc_feature

LOCATION: (1974)..(1974)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (2117)..(2117)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (2133)..(2133)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (2155)..(2335)

OTHER INFORMATION: Alu insertion

FEATURE:

NAME/KEY: misc_feature

LOCATION: (2186)..(2186)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (2191)..(2191)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (2367)..(2367)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (2710)..(2780)

OTHER INFORMATION: L1 fragment insertion

FEATURE:

NAME/KEY: misc_feature

LOCATION: (3391)..(3393)

OTHER INFORMATION: splice acceptor site

FEATURE:

NAME/KEY: misc_feature

LOCATION: (3394)..(3577)

OTHER INFORMATION: Exon 2

FEATURE:

NAME/KEY: misc_feature

LOCATION: (3406)..(3408)

OTHER INFORMATION: Translation initiation codon (ATG)

FEATURE:

NAME/KEY: misc_feature

LOCATION: (3578)..(3583)

OTHER INFORMATION: splice donor site

FEATURE:

NAME/KEY: misc_feature

LOCATION: (4123)..(4123)

OTHER INFORMATION: n = a, c, g, or t.

FEATURE:

NAME/KEY: misc_feature

LOCATION: (4161)..(4163)
OTHER INFORMATION: splice acceptor site
FEATURE:
NAME/KEY: misc feature
LOCATION: (4164)..(4379)
OTHER INFORMATION: Exon 3
FEATURE:
NAME/KEY: misc feature
LOCATION: (4274)..(4276)
OTHER INFORMATION: Translation termination codon (TAA)
FEATURE:
NAME/KEY: polyA signal
LOCATION: (4374)..(4379)
OTHER INFORMATION:
US-10-045-180A-1

Alignment Scores:
Pred. No.: 7.29e-18 Length: 4415
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 14 Gaps: 0

US-10-045-180A-6 (1-31) x US-10-045-180A-1 (1-4415)

Oy 1 ILeCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe 20
Db 4181 ATCCGCTTGCAGAACTCTATCGCATTTTGGAGAACATCTTGGAGACCTGCTTC 4240
Oy 21 ILeuGlyGluArgTyrProIleCysCysTyr 31
Db 4241 ATCCTTGGTGAACGCTACCCCATCTGCTGTAC 4273

RESULT 5

US-09-918-995-18198/c
Sequence 18198, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 18198
LENGTH: 337
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (1)..(337)
OTHER INFORMATION: n = A,T,C or G
US-09-918-995-18198

Alignment Scores:
Pred. No.: 0.00788 Length: 337
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 11 Gaps: 0

US-10-045-180A-6 (1-31) x US-09-918-995-18198 (1-337)

Oy 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe 21
Db 198 TGCATTGCGAATACCGCTGCTATGCGAGAACGTCGTAAGAACTGCTATAC 139

Oy 22 LeuGlyGluArgTyrProIleCysCys 30
Db 138 CAGGGAAGACTCTGGCGATTCTGCTGC 112

RESULT 6

US-09-918-995-19070/c
Sequence 19070, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 19070
LENGTH: 337
TYPE: DNA
ORGANISM: Homo sapiens
US-09-918-995-19070

Alignment Scores:
Pred. No.: 0.00788 Length: 337
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 11 Gaps: 0

US-10-045-180A-6 (1-31) x US-09-918-995-19070 (1-337)

Oy 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe 21
Db 198 TGCATTGCGAATACCGCTGCTATGCGAGAACGTCGTAAGAACTGCTATAC 139
Oy 22 LeuGlyGluArgTyrProIleCysCys 30
Db 138 CAGGGAAGACTCTGGCGATTCTGCTGC 112

RESULT 7

US-09-918-995-18323
Sequence 18323, Application US/09918995
Publication No. US20030073623A1
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
FILE REFERENCE: 20411-756
CURRENT APPLICATION NUMBER: US/09/918, 995
CURRENT FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: US/09/235, 076
PRIOR FILING DATE: 1999-01-20
NUMBER OF SEQ ID NOS: 38054
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 18323
LENGTH: 353
TYPE: DNA
ORGANISM: Homo sapiens
US-09-918-995-18323

Alignment Scores:
Pred. No.: 0.00835 Length: 353
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 11 Gaps: 0

US-10-045-180A-6 (1-31) x US-09-918-995-18323 (1-353)


```
Qy      2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
      147 TGTATTGCAGATACCGACGGTGCATTCGAGGAGAACTGCTATGGAACCTGCATCTAC 206
Db      22 LeuGlyGluArgTyrProIleCysCys 30
      207 CAGGGAAGACTCTGGGCATTCTGCTGC 233

RESULT 8
US-10-252-157-223
; Sequence 223, Application US/10252157
; Publication No. US20030190640A1
; GENERAL INFORMATION:
; APPLICANT: Pearson, Mary
; APPLICANT: Pearson, Cecelia I.
; TITLE OF INVENTION: GENES EXPRESSED IN PROSTATE CANCER
; FILE REFERENCE: PA-0027-1 US/10/252,157
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/295,048
; PRIOR FILING DATE: 2001-05-31
; NUMBER OF SEQ ID NOS: 501
; SOFTWARE: PERL Program
; SEQ ID NO 223
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030190640A1 903338.12
US-10-252-157-223

Alignment Scores:
Pred. No.:      0.0146      Length:      555
Score:          80.00      Matches:      13
Percent Similarity: 55.17%      Conservative: 3
Best Local Similarity: 44.83%      Mismatches: 13
Query Match:    41.88%      Indels:      0
DB:             13          Gaps:         0

US-10-045-180A-6 (1-31) x US-10-252-157-223 (1-555)

Qy      2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
      340 TGTATTGCAGATACCGACGGTGCATTCGAGGAGAACTGCTATGGAACCTGCATCTAC 399
Db      22 LeuGlyGluArgTyrProIleCysCys 30
      400 CAGGGAAGACTCTGGGCATTCTGCTGC 426

RESULT 9
US-09-816-828-16
; Sequence 16, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Zhilwei
; APPLICANT: Zhao, Qiong A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816, 828
; CURRENT FILING DATE: 2001-03-22
```

```
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pc_FL_genes Version 2.0
; SEQ ID NO 16
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(522)
US-09-816-828-16

Alignment Scores:
Pred. No.:      0.0204      Length:      726
Score:          80.00      Matches:      13
Percent Similarity: 55.17%      Conservative: 3
Best Local Similarity: 44.83%      Mismatches: 13
Query Match:    41.88%      Indels:      0
DB:             10          Gaps:         0

US-10-045-180A-6 (1-31) x US-09-816-828-16 (1-726)

Qy      2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
      436 TGTATTGCAGATACCGACGGTGCATTCGAGGAGAACTGCTATGGAACCTGCATCTAC 495
Db      22 LeuGlyGluArgTyrProIleCysCys 30
      496 CAGGGAAGACTCTGGGCATTCTGCTGC 522

RESULT 10
US-09-816-828-16/c
; Sequence 16, Application US/09816828
; Patent No. US20020150898A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Xue, Aidong J.
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Zhilwei
; APPLICANT: Zhao, Qiong A.
; APPLICANT: Zhang, Jie
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and
; FILE REFERENCE: 791CIP2E
; CURRENT APPLICATION NUMBER: US/09/816, 828
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: 09/552,929
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: pc_FL_genes Version 2.0
; SEQ ID NO 16
; LENGTH: 726
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (151)..(522)
US-09-816-828-16

Alignment Scores:
Pred. No.:      0.0204      Length:      726
Score:          80.00      Matches:      13
```


; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(412)
; OTHER INFORMATION: n = A,T,C or G
US-09-736-457-1047

Alignment Scores:
Pred. No.: 0.078 Length: 412
Score: 74.00 Matches: 13
Percent Similarity: 51.72% Conservative: 2
Best Local Similarity: 44.83% Mismatches: 14
Query Match: 38.74% Indels: 0
DB: 10 Gaps: 0

US-10-045-180A-6 (1-31) x US-09-736-457-1047 (1-412)

Qy 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheile 21
Db 254 TGTATTGCCGACCGCGCTGTGTCTACCGGTGAGTCCCTCTCCGGGGTGTGTGAATC 195
Qy 22 LeuGlyGluArgTyrProIleCysCys 30
Db 194 AGTGGCGCGCTNTACAGACTCTGCTGT 168

RESULT 14

US-09-902-941-1047/c
; Sequence 1047, Application US/09902941
; Patent No. US20020172952A1

; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Mannerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1047
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 183, 271, 287, 292, 294, 343
; OTHER INFORMATION: n = A,T,C or G
US-09-902-941-1047

Alignment Scores:
Pred. No.: 0.078 Length: 412
Score: 74.00 Matches: 13
Percent Similarity: 51.72% Conservative: 2
Best Local Similarity: 44.83% Mismatches: 14
Query Match: 38.74% Indels: 0
DB: 10 Gaps: 0

US-10-045-180A-6 (1-31) x US-09-902-941-1047 (1-412)

Qy 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheile 21
Db 254 TGTATTGCCGACCGCGCTGTGTCTACCGGTGAGTCCCTCTCCGGGGTGTGTGAATC 195

Qy 22 LeuGlyGluArgTyrProIleCysCys 30
Db 194 AGTGGCGCGCTNTACAGACTCTGCTGT 168

RESULT 15

US-09-849-626-1047/c
; Sequence 1047, Application US/09849626
; Publication No. US20020197669A1

; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1047
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(412)
; OTHER INFORMATION: n = A,T,C or G
US-09-849-626-1047

Alignment Scores:
Pred. No.: 0.078 Length: 412
Score: 74.00 Matches: 13
Percent Similarity: 51.72% Conservative: 2
Best Local Similarity: 44.83% Mismatches: 14
Query Match: 38.74% Indels: 0
DB: 10 Gaps: 0

US-10-045-180A-6 (1-31) x US-09-849-626-1047 (1-412)

Qy 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheile 21
Db 254 TGTATTGCCGACCGCGCTGTGTCTACCGGTGAGTCCCTCTCCGGGGTGTGTGAATC 195
Qy 22 LeuGlyGluArgTyrProIleCysCys 30
Db 194 AGTGGCGCGCTNTACAGACTCTGCTGT 168

Search completed: December 17, 2003, 21:36:03
Job time : 109.181 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: December 17, 2003, 14:50:33 ; Search time 31 Seconds
(without alignments)
441.383 Million cell updates/sec

Title: US-10-045-180A-6

Perfect score: 191

Sequence: 1 ICHCKVLYCIFGHEHGGTCTILGRYPICCY 31

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool.p/US10045180/runat_17122003_145024_12553/app_query.fasta_1.860
-DB=Issued Patents NA -QFMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10045180 @CGN_1_1_193 @runat_17122003_145024_12553 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DBV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTCUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	191	100.0	453	4	US-09-486-580A-2
2	191	100.0	4415	4	US-09-486-580A-1
3	80	41.9	90	3	US-09-230-180-27
4	80	41.9	108	1	US-08-282-030-43
5	80	41.9	108	1	US-08-282-030-44
6	80	41.9	108	1	US-08-282-030-45
7	80	41.9	108	5	PCT-US95-10219-43
8	80	41.9	108	5	PCT-US95-10219-44
9	80	41.9	108	5	PCT-US95-10219-45
10	80	41.9	109	1	US-08-282-030-61
11	80	41.9	109	5	PCT-US95-10219-61
12	80	41.9	1097	1	US-08-021-608D-5

80	41.9	1097	1	US-08-726-160-5	Sequence 5, Appli
80	41.9	1097	5	PCT-US94-01782-5	Sequence 5, Appli
74	38.7	123	1	US-08-282-030-50	Sequence 50, Appl
74	38.7	123	1	US-08-282-030-62	Sequence 62, Appl
74	38.7	123	5	PCT-US95-10219-50	Sequence 50, Appl
74	38.7	123	5	PCT-US95-10219-62	Sequence 62, Appl
74	38.7	412	4	US-09-702-705-1047	Sequence 1047, Ap
74	38.7	412	4	US-09-736-457-1047	Sequence 1047, Ap
74	38.7	424	1	US-08-158-189-4	Sequence 4, Appli
74	38.7	424	1	US-09-450-972-6	Sequence 6, Appli
74	38.7	2880	1	US-08-158-189-1	Sequence 1, Appli
69.5	36.4	365	1	US-08-486-013-43	Sequence 43, Appl
69.5	36.4	365	2	US-08-482-279-43	Sequence 43, Appl
69.5	36.4	365	2	US-08-342-268-43	Sequence 43, Appl
69.5	36.4	365	3	US-09-015-968-43	Sequence 43, Appl
69.5	36.4	365	4	US-09-397-386-43	Sequence 43, Appl
69	36.1	420	1	US-08-486-013-46	Sequence 46, Appl
69	36.1	420	2	US-08-482-279-46	Sequence 46, Appl
69	36.1	420	2	US-08-342-268-46	Sequence 46, Appl
69	36.1	420	3	US-09-015-968-46	Sequence 46, Appl
69	36.1	420	4	US-09-397-386-46	Sequence 46, Appl
67	35.1	342	1	US-08-486-013-47	Sequence 47, Appl
67	35.1	342	2	US-08-482-279-47	Sequence 47, Appl
67	35.1	342	2	US-08-342-268-47	Sequence 47, Appl
67	35.1	342	3	US-09-015-968-47	Sequence 47, Appl
67	35.1	342	4	US-08-486-013-52	Sequence 52, Appl
67	35.1	388	1	US-08-482-279-52	Sequence 52, Appl
67	35.1	388	2	US-08-342-268-52	Sequence 52, Appl
67	35.1	388	3	US-09-015-968-52	Sequence 52, Appl
67	35.1	388	4	US-09-397-386-52	Sequence 52, Appl
66.5	34.8	452	1	US-08-158-189-6	Sequence 6, Appli
66.5	34.8	452	4	US-09-016-434-1198	Sequence 1198, Ap

ALIGNMENTS

RESULT 1

US-09-486-580A-2

; Sequence 2, Application US/09486580A

; Patent No. 6329340

; GENERAL INFORMATION:

; APPLICANT: GENSET SA

; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC

; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESS: Knobbe, Martens, Olson & Bear

; STREET: 550 West C Street

; CITY: San Diego

; STATE: California

; COUNTRY: USA

; ZIP: 92101

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/486,580A

; FILING DATE: FEBRUARY 25, 2000

; ATTORNEY/AGENT INFORMATION:

; NAME: Hart, Daniel

; REGISTRATION NUMBER: 40,637

; REFERENCE/DOCKET NUMBER: GENSET.064C1

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 453 BASE PAIRS

; TYPE: NUCLEOTIDE

; STRANDEDNESS: DOUBLE

; TOPOLOGY: LINEAR

; MOLECULE TYPE: cDNA

```
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-486-580A-2

Alignment Scores:
Pred. No.: 3,5e-19 Length: 453
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-045-180A-6 (1-31) x US-09-486-580A-2 (1-453)
QY 1 IleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe 20
Db 241 ATCTGCCATTGCAGAGTACTATCTGCAATTTTGGAGAACATCTTGGTGGACCTGCTTC 300
QY 21 IleLeuGlyGluArgTyrProIleCysCysTyr 31
Db 301 ATCTTGGTGAACGCTACCCCAATCTGCTGTAC 333

RESULT 2
US-09-486-580A-1
; Sequence 1, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/POCKET NUMBER: GENSET,064C1
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4415 BASE PAIRS
; TYPE: NUCLEOTIDE
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: DNA
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Exon 1
; LOCATION: 1836..1874
; FEATURE:
; NAME/KEY: Exon 2
; LOCATION: 3394..3577
; FEATURE:
; NAME/KEY: Exon 3
; LOCATION: 4161..4380
; FEATURE:
; NAME/KEY: start CDS
; LOCATION: 3406..3408
; FEATURE:
; NAME/KEY: stop CDS
; LOCATION: 4276..4278
; FEATURE:
; NAME/KEY: polyadenylation site
; LOCATION: 4374..4379
US-09-486-580A-1

Alignment Scores:
Pred. No.: 6,14e-18 Length: 4415
Score: 191.00 Matches: 31
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-10-045-180A-6 (1-31) x US-09-486-580A-1 (1-4415)
QY 1 IleCysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe 20
Db 4181 ATCTGCCATTGCAGAGTACTATCTGCAATTTTGGAGAACATCTTGGTGGACCTGCTTC 4240
QY 21 IleLeuGlyGluArgTyrProIleCysCysTyr 31
Db 4241 ATCTTGGTGAACGCTACCCCAATCTGCTGTAC 4273

RESULT 3
US-09-230-180-27
; Sequence 27, Application US/09230180
; Patent No. 6183992
; GENERAL INFORMATION:
; APPLICANT: Kim, Sun-Chang
; APPLICANT: Lee, Jae Hyun
; APPLICANT: Kang, Min Hyun
; APPLICANT: Kim, Jeong Hyun
; APPLICANT: Hong, Seung-Suh
; APPLICANT: Lee, Hyun-Soo
; APPLICANT: Samsung Genex Corporation
; APPLICANT: Korea Advanced Institute of Science and Technology
; TITLE OF INVENTION: METHOD FOR MASS PRODUCTION OF
; TITLE OF INVENTION: ANTIMICROBIAL PEPTIDE
; FILE REFERENCE: 6181/0F135
; CURRENT APPLICATION NUMBER: US/09/230,180
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: PCT/KR98/00132
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: KR 13372/1998
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: KR 21312/1997
; PRIOR FILING DATE: 1997-05-28
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 90
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: DNA sequence deduced from HNP-I peptide sequence
; OTHER INFORMATION: based on codon usage of E. coli
US-09-230-180-27

Alignment Scores:
Pred. No.: 0.00058 Length: 90
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 3 Gaps: 0

US-10-045-180A-6 (1-31) x US-09-230-180-27 (1-90)
QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe 21
Db 4 TGCTACTGCGGTATCCCGCGTGCATCGCGGTGAGCGGTGCTGTACGCTGCTAC 63
```

Qy 22 LeuGlyGluArgTyrProIleCysCys 30
 Db 64 CAGGGTCGTCGTGGCGCTTCGTGTC 90

RESULT 4

US-08-282-030-43
 ; Sequence 43, Application US/08282030
 ; Patent No. 5589364
 ; GENERAL INFORMATION:
 ; APPLICANT: Williams, Jon I.
 ; APPLICANT: Pierce, James C.
 ; APPLICANT: Anderson, Mark G.
 ; APPLICANT: Kari, Prasad
 ; TITLE OF INVENTION: Recombinant Production of Biologically
 ; TITLE OF INVENTION: Active Peptides and Proteins
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315

COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION NUMBER: US/08/282,030
 ; FILING DATE: 29-JUL-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fordis, Jean B.
 ; REGISTRATION NUMBER: 32,984
 ; REFERENCE/DOCKET NUMBER: 05387.0001-00000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-408-4400
 ; TELEFAX: 202-408-4400
 ; INFORMATION FOR SEQ ID NO: 43:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 108 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-282-030-43

Alignment Scores:
 Pred. No.: 0.00073 Length: 108
 Score: 80.00 Matches: 13
 Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: 1 Gaps: 0

US-10-045-180A-6 (1-31) x US-08-282-030-43 (1-108)

Qy 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
 Db 12 TGTACTGCGTATTCGGCATGCGATCGCAGCGAGCGTGTGCTACTGTATTATAC 71
 Qy 22 LeuGlyGluArgTyrProIleCysCys 30
 Db 72 CAGGGTCGTCGTGGCGCATTCGTGTC 98

RESULT 5

US-08-282-030-44/c
 ; Sequence 44, Application US/08282030
 ; Patent No. 5589364
 ; GENERAL INFORMATION:
 ; APPLICANT: Williams, Jon I.

; APPLICANT: Pierce, James C.
 ; APPLICANT: Anderson, Mark G.
 ; APPLICANT: Kari, Prasad
 ; TITLE OF INVENTION: Recombinant Production of Biologically
 ; TITLE OF INVENTION: Active Peptides and Proteins
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3315
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION NUMBER: US/08/282,030
 ; FILING DATE: 29-JUL-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fordis, Jean B.
 ; REGISTRATION NUMBER: 32,984
 ; REFERENCE/DOCKET NUMBER: 05387.0001-00000
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-408-4400
 ; TELEFAX: 202-408-4400
 ; INFORMATION FOR SEQ ID NO: 44:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 108 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; US-08-282-030-44

Alignment Scores:
 Pred. No.: 0.00073 Length: 108
 Score: 80.00 Matches: 13
 Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: 1 Gaps: 0

US-10-045-180A-6 (1-31) x US-08-282-030-44 (1-108)

Qy 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
 Db 101 TGTACTGCGTATTCGGCATGCGATCGCAGCGAGCGTGTGCTACTGTATTATAC 42
 Qy 22 LeuGlyGluArgTyrProIleCysCys 30
 Db 41 CAGGGTCGTCGTGGCGCATTCGTGTC 15

RESULT 6

US-08-282-030-45
 ; Sequence 45, Application US/08282030
 ; Patent No. 5589364
 ; GENERAL INFORMATION:

; APPLICANT: Williams, Jon I.
 ; APPLICANT: Pierce, James C.
 ; APPLICANT: Anderson, Mark G.
 ; APPLICANT: Kari, Prasad
 ; TITLE OF INVENTION: Recombinant Production of Biologically
 ; TITLE OF INVENTION: Active Peptides and Proteins
 ; NUMBER OF SEQUENCES: 62
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
 ; STREET: 1300 I Street, N.W.
 ; CITY: Washington

```
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,030
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-282-030-45

Alignment Scores:
Pred. No.: 0.00073 Length: 108
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-6 (1-31) x US-08-282-030-45 (1-108)
QY 2 CysHisCysArgValLeuTyrcysIlePheGlyGluHisLeuGlyThrCysPheIle 21
Db 12 TGTACTCGCGTATTCGGCATGCATCGCAGCGCGGTCGTATGTTATTAC 71
QY 22 LeuGlyGluArgTyrcysIleCysCys 30
Db 72 CAGGTCGTCTGTGGCAATCTGTGTC 98

RESULT 7
PCT-US95-10219-43
; Sequence 43, Application PC/TUS9510219
; GENERAL INFORMATION:
; APPLICANT: Williams, Jon I.
; APPLICANT: Pierce, James C.
; APPLICANT: Anderson, Mark G.
; APPLICANT: Kari, Praead
; TITLE OF INVENTION: Recombinant Production of Biologically
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10219
; FILING DATE:
; CLASSIFICATION:
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/282,030
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 43:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; PCT-US95-10219-43

Alignment Scores:
Pred. No.: 0.00073 Length: 108
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 5 Gaps: 0

US-10-045-180A-6 (1-31) x PCT-US95-10219-43 (1-108)
QY 2 CysHisCysArgValLeuTyrcysIlePheGlyGluHisLeuGlyThrCysPheIle 21
Db 12 TGTACTCGCGTATTCGGCATGCATCGCAGCGCGGTCGTATGTTATTAC 71
QY 22 LeuGlyGluArgTyrcysIleCysCys 30
Db 72 CAGGTCGTCTGTGGCAATCTGTGTC 98

RESULT 8
PCT-US95-10219-44/c
; Sequence 44, Application PC/TUS9510219
; GENERAL INFORMATION:
; APPLICANT: Williams, Jon I.
; APPLICANT: Pierce, James C.
; APPLICANT: Anderson, Mark G.
; APPLICANT: Kari, Praead
; TITLE OF INVENTION: Recombinant Production of Biologically
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10219
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/282,030
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
```

```

; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US95-10219-44

Alignment Scores:
Pred. No.: 0.00073 Length: 108
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservatives: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 5 Gaps: 0

US-10-045-180A-6 (1-31) x PCT-US95-10219-44 (1-108)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe11
Db 101 TGTACTGCGTATTCGGCATGTCGAGCGGCGTATGCTGTTGTTTAC 42
QY 22 LeuGlyGluArgTyrProIleCysCys 30
Db 41 CAGGGTCGTCGTGGGCATTCTGTTC 15

RESULT 9
PCT-US95-10219-45
; Sequence 45, Application PC/TUS9510219
; GENERAL INFORMATION:
; APPLICANT: Williams, Jon I.
; APPLICANT: Pierce, James C.
; APPLICANT: Anderson, Mark G.
; APPLICANT: Kari, Prasad
; TITLE OF INVENTION: Recombinant Production of Biologically
; TITLE OF INVENTION: Active Peptides and Proteins
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10219
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/282,030
; FILING DATE: 29-JUL-1994
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 108 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
PCT-US95-10219-45

```

```

Alignment Scores:
Pred. No.: 0.00073 Length: 108
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservatives: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 5 Gaps: 0

US-10-045-180A-6 (1-31) x PCT-US95-10219-45 (1-108)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe11
Db 12 TGTACTGCGTATTCGGCATGTCGAGCGGCGTATGCTGTTTATTTAC 71
QY 22 LeuGlyGluArgTyrProIleCysCys 30
Db 72 CAGGGTCGTCGTGGGCATTCTGTTC 98

RESULT 10
US-08-282-030-61/c
; Sequence 61, Application US/08282030
; Patent No. 5589364
; GENERAL INFORMATION:
; APPLICANT: Williams, Jon I.
; APPLICANT: Pierce, James C.
; APPLICANT: Anderson, Mark G.
; APPLICANT: Kari, Prasad
; TITLE OF INVENTION: Recombinant Production of Biologically
; TITLE OF INVENTION: Active Peptides and Proteins
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,030
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-282-030-61

Alignment Scores:
Pred. No.: 0.000738 Length: 109
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservatives: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-6 (1-31) x US-08-282-030-61 (1-109)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPhe11

```


Db 102 TGTACTGCGTATTCGCGCATCGACGGAGCGCTCGCTATGGTACTGTGATTATTC 43
QY 22 LeuGlyGluArgTyrProIleCysCys 30
Db 42 CAGGGTCGTCTGTGGGCATTCTGTTC 16

RESULT 11

PCT-US95-10219-61/c
; Sequence 61, Application PC/TUS9510219
; GENERAL INFORMATION:
; APPLICANT: Williams, Jon I.
; APPLICANT: Pierce, James C.
; APPLICANT: Anderson, Mark G.
; APPLICANT: Kari, Prasad
; TITLE OF INVENTION: Recombinant Production of Biologically
; TITLE OF INVENTION: Active Peptides and Proteins
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10219
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/282,030
; FILING DATE: 29-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 109 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

PCT-US95-10219-61
Alignment Scores:
Pred. No.: 0.000738 Length: 109
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 5 Gaps: 0

US-10-045-180A-6 (1-31) x PCT-US95-10219-61 (1-109)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyThrCysPheIle 21
Db 102 TGTACTGCGTATTCGCGCATCGACGGAGCGCTCGCTATGGTACTGTGATTATTC 43
QY 22 LeuGlyGluArgTyrProIleCysCys 30
Db 42 CAGGGTCGTCTGTGGGCATTCTGTTC 16

RESULT 12

US-08-021-608D-5/c

; Sequence 5, Application US/08021608D
; Patent No. 5580760
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/021,608D
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26,728
; REFERENCE/DOCKET NUMBER: 2026-4063
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1097
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: No
; ANTI-SENSE: YES
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; US-08-021-608D-5

Alignment Scores:
Pred. No.: 0.0135 Length: 1097
Score: 80.00 Matches: 13
Percent Similarity: 55.17% Conservative: 3
Best Local Similarity: 44.83% Mismatches: 13
Query Match: 41.88% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-6 (1-31) x US-08-021-608D-5 (1-1097)

QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyThrCysPheIle 21
Db 891 TGTATTGCAATACCGCGTGCATTGCGAGAGACGTCGCTATGGAACCTGCATCTAC 832
QY 22 LeuGlyGluArgTyrProIleCysCys 30
Db 831 CAGGGAAGACTCTGGGCATTCTGTTC 805

RESULT 13

US-08-726-160-5/c

; Sequence 5, Application US/08726160
; Patent No. 5734016
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,

APPLICANT: ROBERT C., AND AVIGAN, MARK I.
 TITLE OF INVENTION: NOVEL FUSE BINDING
 NUMBER OF SEQUENCES: 24
 CORRESPONDENCE ADDRESS:

ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154

COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/726,160
 FILING DATE: 04-OCT-1996

CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/021,608
 FILING DATE: 22-FEB-1993
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:

NAME: WILLIAM S. FEILER
 REGISTRATION NUMBER: 26,728
 REFERENCE/DOCKET NUMBER: 2026-4063US1
 TELEPHONE: (212) 751-6849
 TELEFAX: (212) 751-6849

INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1097
 TYPE: Nucleic acid
 STRANDEDNESS: Double

TOPOLOGY: Unknown
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: No
 ANTI-SENSE: YES
 ORIGINAL SOURCE:
 ORGANISM: Human
 CELL LINE: HL60

US-08-726-160-5

Alignment Scores:
 Pred. No.: 0.0135 Length: 1097
 Score: 80.00 Matches: 13
 Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: 1 Gaps: 0

US-10-045-180A-6 (1-31) x US-08-726-160-5 (1-1097)

Qy 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
 Db 891 TGCATTGCGAGATACCAACGCGTGCATTGCGAGAGACGTCGCTATGGAACCTGCATCTAC 832
 Qy 22 LeuGlyGluArgTyrProIleCysCys 30
 Db 831 CAGGAAGACTCTGGGCATTCTGCTGC 805

RESULT 14

PCT-US94-01782-5/c
 Sequence 5, Application PC/TUS9401782

GENERAL INFORMATION:
 APPLICANT: THE GOVERNMENT OF THE UNITED STATES
 APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
 APPLICANT: HEALTH AND HUMAN SERVICES
 TITLE OF INVENTION: NOVEL FUSE BINDING
 TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
 NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US94/01782
 FILING DATE: 22-FEB-1994
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: U.S. 08/021,608
 FILING DATE: 22-FEB-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: WILLIAM S. FEILER
 REGISTRATION NUMBER: 26,728
 REFERENCE/DOCKET NUMBER: 2026-4063PCT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 751-6849
 TELEFAX: (212) 751-6849
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1097
 TYPE: Nucleic acid
 STRANDEDNESS: Double
 TOPOLOGY: Unknown
 MOLECULE TYPE: cDNA
 HYPOTHETICAL: No
 ANTI-SENSE: YES
 ORIGINAL SOURCE:
 ORGANISM: Human
 STRAIN:
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE:
 CELL LINE: HL60
 ORGANELLE:
 POSITION IN GENOME:
 CHROMOSOME/SEGMENT:
 MAP POSITION:
 UNITS:
 FEATURE:
 NAME/KEY:
 LOCATION:
 IDENTIFICATION METHOD:
 OTHER INFORMATION:

PCT-US94-01782-5

Alignment Scores:
 Pred. No.: 0.0135 Length: 1097
 Score: 80.00 Matches: 13
 Percent Similarity: 55.17% Conservative: 3
 Best Local Similarity: 44.83% Mismatches: 13
 Query Match: 41.88% Indels: 0
 DB: 5 Gaps: 0

US-10-045-180A-6 (1-31) x PCT-US94-01782-5 (1-1097)

Qy 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
 Db 891 TGCATTGCGAGATACCAACGCGTGCATTGCGAGAGACGTCGCTATGGAACCTGCATCTAC 832
 Qy 22 LeuGlyGluArgTyrProIleCysCys 30
 Db 831 CAGGAAGACTCTGGGCATTCTGCTGC 805

Thu Dec 18 06:09:44 2003

RESULT 15
US-08-282-030-50
; Sequence 50, Application US/08282030
; Patent No. 5589364
; GENERAL INFORMATION:
; APPLICANT: Williams, Jon I.
; APPLICANT: Pierce, James C.
; APPLICANT: Anderson, Mark G.
; APPLICANT: Kari, Prasad
; TITLE OF INVENTION: Recombinant Production of Biologically
; TITLE OF INVENTION: Active Peptides and Proteins
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/282,030
; FILING DATE: 29-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 05387.0001-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 123 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-282-030-50

Alignment Scores:
Pred. No.: 0.00638 Length: 123
Score: 74.00 Matches: 13
Percent Similarity: 51.72% Conservative: 2
Best Local Similarity: 44.83% Mismatches: 14
Query Match: 38.74% Indels: 0
DB: 1 Gaps: 0

US-10-045-180A-6 (1-31) x US-08-282-030-50 (1-123)
QY 2 CysHisCysArgValLeuTyrCysIlePheGlyGluHisLeuGlyGlyThrCysPheIle 21
Db 24 TGCCTACTGTCGACTGGTGTGTGCAACGCGTGAAGCCTGAGCGCGTCTGTGGAATC 83
QY 22 LeuGlyGluArgTyrProIleCysCys 30
Db 84 TCCGGTCGTCTGTATCGCCTGTGTTC 110

Search completed: December 17, 2003, 14:54:03
Job time : 32 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 : Search time 20.304 Seconds
(without alignments)
445.225 Million cell updates/sec

Title: US-10-045-180A-3

Perfect score: 94

Sequence: 1 MRTITLSAFLLVALQAAE.....GEHLGGTCFILGERYPICCY 94

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

1: Pir1:*

2: Pir2:*

3: Pir3:*

4: Pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	8.5	94	1 A40499	defensin alpha-1 p
2	8	8.5	94	1 B40499	defensin alpha-3 p
3	8	8.5	128	2 T15912	hypothetical prote
4	7	7.4	87	2 H64521	hypothetical prote
5	7	7.4	87	2 D71984	hypothetical prote
6	7	7.4	94	1 A44454	defensin alpha-5 p
7	7	7.4	95	1 WTRBM1	macrophage antibio
8	7	7.4	95	1 WTRBM2	defensin CS-4 prec
9	7	7.4	100	1 S27016	defensin alpha-6 p
10	7	7.4	183	2 D81396	MAF homolog Cj0507
11	7	7.4	196	2 A48150	hibernation-relate
12	7	7.4	235	2 H96023	probable phosphona
13	7	7.4	286	2 A28333	carcinoembryonic a
14	7	7.4	321	2 JH0395	biliary glycoprote
15	7	7.4	328	1 S16300	UDPglucose 4-epime
16	7	7.4	328	2 A95975	probable UDPglucos
17	7	7.4	328	2 S60256	TGF-beta receptor
18	7	7.4	328	2 H84900	hypothetical prote
19	7	7.4	339	2 T34925	ABC transporter in
20	7	7.4	347	2 B71347	probable catabolit
21	7	7.4	349	2 A34815	carcinoembryonic a
22	7	7.4	351	2 JH0396	biliary glycoprote
23	7	7.4	412	2 G81581	tyrosyl-tRNA synth
24	7	7.4	412	2 H71560	tyrosine-tRNA liga
25	7	7.4	417	2 JH0394	biliary glycoprote
26	7	7.4	422	2 G86535	tyrosyl tRNA synth
27	7	7.4	422	2 D72088	tyrosine-tRNA liga
28	7	7.4	436	2 T36432	probable transmem
29	7	7.4	464	2 C30127	transmembrane carc

30	7	7.4	526	1 A32164	biliary glycoprote
31	7	7.4	628	2 A87596	hypothetical prote
32	7	7.4	702	2 A36319	carcinoembryonic a
33	7	7.4	1450	2 JG6139	cystic fibrosis tr
34	7	7.4	1476	1 A39901	cystic fibrosis tr
35	7	7.4	1476	1 A40303	cystic fibrosis tr
36	7	7.4	1480	1 DVHUCF	cystic fibrosis tr
37	7	7.4	1485	2 S23756	CFTF protein - Afr
38	6	6.4	31	2 T09121	hypothetical prote
39	6	6.4	57	2 D86108	hypothetical prote
40	6	6.4	85	2 JC2290	hypothetical prote
41	6	6.4	85	2 AD1537	B. subtilis protei
42	6	6.4	89	2 S25602	ubiquinol-cytochro
43	6	6.4	90	2 T28989	hypothetical prote
44	6	6.4	97	1 A47365	defensin alpha-4 p
45	6	6.4	103	1 IKECV	colicin V - Escher
46	6	6.4	106	2 A88650	protein C09G12.5 [
47	6	6.4	115	1 G64304	hypothetical prote
48	6	6.4	117	2 F69887	hypothetical prote
49	6	6.4	124	2 S40099	hypothetical prote
50	6	6.4	124	2 H82117	hypothetical prote
51	6	6.4	131	1 D69827	hypothetical prote
52	6	6.4	131	2 AG3198	hypothetical prote
53	6	6.4	134	2 JC5116	interleukin-5 prec
54	6	6.4	137	2 A70913	hypothetical prote
55	6	6.4	142	2 D72340	hypothetical prote
56	6	6.4	144	2 T18016	hypothetical prote
57	6	6.4	157	2 A31084	Gdx protein - huma
58	6	6.4	161	2 F70006	cysteine dioxygena
59	6	6.4	165	2 C69065	hypothetical prote
60	6	6.4	168	2 T17592	hypothetical prote
61	6	6.4	168	2 AG2986	conserved hypothet
62	6	6.4	168	2 H98296	hypothetical prote
63	6	6.4	172	2 H81375	probable integral
64	6	6.4	174	2 A49181	alpha B-crystallin
65	6	6.4	174	2 S58758	alpha-Crystallin C
66	6	6.4	175	2 A70952	hypothetical prote
67	6	6.4	176	2 T21847	hypothetical prote
68	6	6.4	179	2 E46660	peptidoglycan asso
69	6	6.4	179	2 C89467	conserved hypothet
70	6	6.4	180	2 I48129	Xel69 (escapes X-1
71	6	6.4	180	2 I84689	escapes X-inactiva
72	6	6.4	181	2 AF2468	50S ribosomal prot
73	6	6.4	183	2 AF0904	probable exported
74	6	6.4	184	2 T40305	hypothetical prote
75	6	6.4	186	2 T11384	Arfase subunit 6 -
76	6	6.4	187	2 D85562	hypothetical prote
77	6	6.4	187	2 H64796	ybeM protein - Esc
78	6	6.4	187	2 T51719	hypothetical prote
79	6	6.4	189	2 F82969	hypothetical prote
80	6	6.4	191	2 G64017	hypothetical prote
81	6	6.4	195	2 D81059	hypothetical prote
82	6	6.4	195	2 B85904	hypothetical prote
83	6	6.4	195	2 A65036	hypothetical 21.2
84	6	6.4	197	2 A10115	conserved hypothet
85	6	6.4	197	2 T26829	hypothetical prote
86	6	6.4	197	2 C98809	protein Y43CB.1 [
87	6	6.4	205	2 C84509	hypothetical prote
88	6	6.4	207	2 T38961	hypothetical prote
89	6	6.4	212	2 AG0854	probable sugar ald
90	6	6.4	213	2 F84581	copia-like retrosl
91	6	6.4	219	2 T51607	glutathione transf
92	6	6.4	225	2 C83316	NADH dehydrogenase
93	6	6.4	225	2 AC1896	hypothetical prote
94	6	6.4	226	2 D49591	membrane protein M
95	6	6.4	226	2 S37434	membrane glycoprot
96	6	6.4	226	2 C90877	probable tail asse
97	6	6.4	226	2 D90908	probable tail asse
98	6	6.4	226	2 F58842	conserved hypothet
99	6	6.4	227	2 E69079	conserved hypothet
100	6	6.4	231	2 T12078	ribonuclease (EC 3

ALIGNMENTS

RESULT 1

A40499
defensin alpha-1 precursor, myeloid-related [validated] - human
N;Alternate names: defensin HNP-1; myeloid leukocyte protein mrs; myeloid-related protein
N;Contains: defensin HNP-1; defensin HNP-3
C;Species: Homo sapiens (man)
C;Date: 24-Jan-1992 #sequence revision 24-Jan-1992 #text change 08-Dec-2000
C;Accession: S32499; S07161; A40499; A30658; B30658; A32958; S65412; S65414
R;Linzeimer, R.; Michaelson, D.; Liu, L.; Ganz, T.
FEBs Lett. 321, 267-273, 1993
A;Title: The structure of neutrophil defensin genes.
A;Reference number: S32499; MUID:93238968; PMID:8477861
A;Accession: S32499
A;Molecule type: DNA
A;Residues: 1-94 <LIN>
R;Cross-references: GB:112690; NID:G292362; PIDN:AAA52302.1; PID:G292363
R;Mars, W.M.; van Tuinen, P.; Drabkin, H.A.; White, J.W.; Saunders, G.F.
Blood 71, 1713-1719, 1988
A;Title: A myeloid-related sequence that localizes to human chromosome 8q21.1-22.
A;Reference number: S07161; MUID:88222410; PMID:3370315
A;Accession: S07161
A;Molecule type: mRNA
A;Residues: 1-94 <MAR>
R;Cross-references: GB:M26602; NID:G181528; PIDN:AAA52303.1; PID:G181529
R;Daher, K.A.; Lehrer, R.I.; Ganz, T.; Kronenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 85, 7327-7331, 1988
A;Title: Isolation and characterization of human defensin cDNA clones.
A;Reference number: A40499; MUID:89017191; PMID:3174637
A;Accession: A40499
A;Molecule type: mRNA
A;Residues: 1-94 <DAH>
R;Cross-references: GB:M21130; GB:M22160; NID:G181526; PIDN:AAA52302.1; PID:G181527
R;Seisted, M.E.; Harwig, S.S.L.; Ganz, T.; Schilling, J.W.; Lehrer, R.I.
J. Clin. Invest. 76, 1436-1439, 1985
A;Title: Primary structures of three human neutrophil defensins.
A;Reference number: A30658; MUID:86034582; PMID:4056036
A;Accession: A30658
A;Molecule type: protein
A;Residues: 65-94 <SEL1>
A;Experimental source: HNP-1
A;Accession: B30658
A;Molecule type: protein
A;Residues: 66-94 <SEL2>
A;Experimental source: HNP-2
R;Seisted, M.E.; Harwig, S.S.L.
J. Biol. Chem. 264, 4003-4007, 1989
A;Title: Determination of the disulfide array in the human defensin HNP-2. A covalently
A;Reference number: A32958; MUID:89139465; PMID:2917986
A;Accession: A32958
A;Molecule type: protein
A;Residues: 66-94 <SEL>
R;Frohman, M.; Gunne, H.; Bergman, A.C.; Agerberth, B.; Bergman, T.; Boman, A.; Liden, S.;
Eur. J. Biochem. 237, 86-92, 1996
A;Title: Biochemical and antibacterial analysis of human wound and blister fluid.
A;Reference number: S65409; MUID:96203912; PMID:8620898
A;Accession: S65412
A;Molecule type: protein
A;Residues: 65, 'X', 67, 'X', 69-72, 'X', 74-82 <FRO1>
A;Accession: S65414
A;Molecule type: protein
A;Residues: 'X', 67, 'X', 69-72, 'X', 74-82, 84 <FRO2>
R;Pardi, A.; Zhang, X.L.; Seisted, M.E.; Skaliky, J.J.; Yip, P.F.
Biochemistry 31, 11357-11364, 1992
A;Title: NMR studies of defensin antimicrobial peptides. 2. Three-dimensional structures
A;Reference number: A44209; MUID:93075734; PMID:1445873
A;Contents: annotation; conformation and disulfide bond assignments by (1)H-NMR
R;Wilson, C.L.; Ouellette, A.J.; Satchell, D.P.; Ayabe, T.; Lopez-Boado, Y.S.; Stratman,
Science 286, 113-117, 1999
A;Title: Regulation of intestinal alpha-defensin activation by the metalloproteinase mat
A;Reference number: A59079; MUID:99438247; PMID:10506557

A;Contents: annotation; activation by matrilysin
C;Comment: The propeptide is cleaved by matrilysin (EC 3.4.24.23) (see PIR:KCHUM) .
C;Genetics:
A;Gene: GDB:DEF1; DEF1; MRS; HNP-1
A;Cross-references: GDB:120745; OMIM:125220
A;Map position: 8p23-8p22
A;Introns: 59/1
C;Complex: homodimer
C;Superfamily: mammalian defensin
C;Keywords: antibiotic; homodimer
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-64/Domain: propeptide #status predicted <PRO>
F;65-94/Product: defensin HNP-1 #status experimental <MAT1>
F;66-94/Product: defensin HNP-2 #status experimental <MAT2>
F;66-94,68-83,73-93/Disulfide bonds: #status experimental

Query Match 8.5%; Score 8; DB 1; Length 94;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 19 AEPLOARA 26

Db 19 AEPLOARA 26

RESULT 2

B40499
defensin alpha-3 precursor, neutrophil-specific [validated] - human
N;Alternate names: defensin HNP-3; neutrophil peptide 3 (HNP-3)
N;Contains: defensin HNP-2; defensin HNP-3
C;Species: Homo sapiens (man)
C;Date: 24-Jan-1992 #sequence revision 03-Nov-1995 #text change 15-Sep-2000
C;Accession: S32500; S08005; B40499; S65413; A32958; C30658
R;Linzeimer, R.; Michaelson, D.; Liu, L.; Ganz, T.
FEBs Lett. 321, 267-273, 1993
A;Title: The structure of neutrophil defensin genes.
A;Reference number: S32499; MUID:93238968; PMID:8477861
A;Accession: S32500
A;Molecule type: DNA
A;Residues: 1-94 <LIN>
R;Wiedemann, L.M.; Francis, G.E.; Lamb, R.F.; Burns, J.H.; Winnie, J.N.; MacKenzie, E.D.
submitted to the EMBL Data Library, November 1988
A;Reference number: S08005
A;Accession: S08005
A;Molecule type: mRNA
A;Residues: 1-94 <WIE>
R;Experimental source: HNP-3 expressed in promyelocytes and myelocytes
R;Daher, K.A.; Lehrer, R.I.; Ganz, T.; Kronenberg, M.
Proc. Natl. Acad. Sci. U.S.A. 85, 7327-7331, 1988
A;Title: Isolation and characterization of human defensin cDNA clones.
A;Reference number: A40499; MUID:89017191; PMID:3174637
A;Accession: B40499
A;Molecule type: mRNA
A;Residues: 30-94 <DAH>
R;Cross-references: GB:M21131; NID:G181530; GB:M23281; NID:G181534; PIDN:AAA52304.1; PID:
A;Experimental source: clone HNP-3
A;Note: the sequence in GenBank entry HUMDEF3A, release 111.0, is translated in an inco
R;Frohman, M.; Gunne, H.; Bergman, A.C.; Agerberth, B.; Bergman, T.; Boman, A.; Liden, S.
Eur. J. Biochem. 237, 86-92, 1996
A;Title: Biochemical and antibacterial analysis of human wound and blister fluid.
A;Reference number: S65409; MUID:96203912; PMID:8620898
A;Accession: S65413
A;Molecule type: protein
A;Residues: 65, 'X', 67, 'X', 69-72, 'X', 74-82 <FRO>
R;Seisted, M.E.; Harwig, S.S.L.
J. Biol. Chem. 264, 4003-4007, 1989
A;Title: Determination of the disulfide array in the human defensin HNP-2. A covalently
A;Reference number: A32958; MUID:89139465; PMID:2917986
A;Accession: A32958
A;Molecule type: protein

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 ; Search time 15.792 Seconds
(without alignments)
279.921 Million cell updates/sec

Title: US-10-045-180A-3

Perfect score: 94

Sequence: 1 MRTLTLALFLVALQAAAE.....GEHLGGTCFLGERYPICCY 94

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	8	8.5	94	1	DEF1_HUMAN
2	8	8.5	94	1	DEF3_HUMAN
3	8	8.5	649	1	S281_HUMAN
4	7	7.4	87	1	DEFA_RAT
5	7	7.4	93	1	DEF4_RAT
6	7	7.4	94	1	DEF1_RAT
7	7	7.4	94	1	DEF2_RAT
8	7	7.4	94	1	DEF5_HUMAN
9	7	7.4	95	1	DEF3_RABIT
10	7	7.4	95	1	DEF4_RABIT
11	7	7.4	100	1	DEF6_HUMAN
12	7	7.4	183	1	PSAF_PROMA
13	7	7.4	183	1	Y507_CAMJE
14	7	7.4	186	1	HP20_TAMSI
15	7	7.4	327	1	EXOB_RHILT
16	7	7.4	328	1	EXOB_RHIME
17	7	7.4	328	1	IF32_ARATH
18	7	7.4	349	1	CEA8_HUMAN
19	7	7.4	412	1	SYI_CHLPN
20	7	7.4	412	1	SYI_CHLTP
21	7	7.4	526	1	CEA1_HUMAN
22	7	7.4	702	1	CEA5_HUMAN
23	7	7.4	1450	1	CFTFR_RABIT
24	7	7.4	1476	1	CFTFR_MOUSE
25	7	7.4	1480	1	CFTFR_HUMAN
26	7	7.4	1485	1	CFTFR_XENLA
27	6	6.4	31	1	PETL_NEPOL
28	6	6.4	31	1	PETL_POPDE
29	6	6.4	52	1	CRAB_COLLI
30	6	6.4	52	1	CRAB_TURME
31	6	6.4	57	1	YUDO_ECOLI
32	6	6.4	67	1	Y12B_BPT4
33	6	6.4	85	1	Y522_PSBAE
					P59665 homo sapien
					P59666 homo sapien
					O00337 homo sapien
					O62713 rattus norv
					O62714 rattus norv
					O62716 rattus norv
					O62715 rattus norv
					O01523 homo sapien
					P01376 oryctolagus
					P01377 oryctolagus
					Q01524 homo sapien
					Q9x714 prochloroco
					Q9pi04 campylobact
					Q06575 tamias sibi
					Q59745 rhizobium l
					P26503 rhizobium m
					Q38884 arabidopsis
					P31997 homo sapien
					Q92812 chlamydia p
					O84085 chlamydia t
					P13688 homo sapien
					Q00554 oryctolagus
					P26361 mus musculus
					P13569 homo sapien
					P26363 xenopus lae
					Q9tky9 nephroselmi
					O20272 populus del
					O12987 columba liv
					O12995 turdus meru
					P58038 escherichia
					P39494 bacteriophag
					Q51483 pseudomonas

RESULT 1

ALIGNMENTS

34	6	6.4	97	1	DEF4_HUMAN
35	6	6.4	103	1	CRAV_ECOLI
36	6	6.4	115	1	Y039_METJA
37	6	6.4	131	1	CRB2_BACSU
38	6	6.4	132	1	IL5_SHEEP
39	6	6.4	134	1	IL5_BOVIN
40	6	6.4	134	1	IL5_CERTO
41	6	6.4	134	1	IL5_HORSE
42	6	6.4	135	1	IL5_CAVPO
43	6	6.4	137	1	Y045_MYCTU
44	6	6.4	153	1	VNSC_TPMV
45	6	6.4	157	1	UBLG_HUMAN
46	6	6.4	174	1	CRAB_ANAPL
47	6	6.4	174	1	CRAB_CHICK
48	6	6.4	180	1	SMCX_CRIGR
49	6	6.4	181	1	BY55_HUMAN
50	6	6.4	181	1	RL10_ANASP
51	6	6.4	185	1	BY55_MOUSE
52	6	6.4	191	1	Y977_HAEIN
53	6	6.4	195	1	YFIK_ECOLI
54	6	6.4	197	1	HAM1_YERPE
55	6	6.4	205	1	YG77_METJA
56	6	6.4	210	1	SEBP_SCHPO
57	6	6.4	210	1	PSBP_PORPU
58	6	6.4	233	1	YJT5_YEAST
59	6	6.4	253	1	EBSD_ENTFA
60	6	6.4	254	1	AROD_NEIMA
61	6	6.4	254	1	AROD_NEIMB
62	6	6.4	258	1	CTR2_ANOGA
63	6	6.4	259	1	CTR1_ANOGA
64	6	6.4	261	1	COX3_PAPHA
65	6	6.4	262	1	VGL1_CVPR8
66	6	6.4	262	1	VME1_CVCAI
67	6	6.4	262	1	VME1_CVPPS
68	6	6.4	262	1	VME1_CVPPU
69	6	6.4	262	1	VME1_CVPRM
70	6	6.4	262	1	VME1_FIPV
71	6	6.4	262	1	YBEM_ECO57
72	6	6.4	262	1	YBEM_ECOLI
73	6	6.4	269	1	COX3_MAGGR
74	6	6.4	277	1	CYST_ECOLI
75	6	6.4	291	1	SH1D_PIG
76	6	6.4	294	1	BLAC_CITDI
77	6	6.4	294	1	MCSR_PIG
78	6	6.4	296	1	YFCH_HABIN
79	6	6.4	302	1	DAPA_XYLFA
80	6	6.4	303	1	T2RD_HUMAN
81	6	6.4	306	1	YC64_AQUAE
82	6	6.4	309	1	TR43_HUMAN
83	6	6.4	309	1	TR52_HUMAN
84	6	6.4	316	1	T2BB_BACSU
85	6	6.4	332	1	Y4RC_RHISN
86	6	6.4	336	1	XYLZ_PSEPU
87	6	6.4	338	1	Y941_METJA
88	6	6.4	342	1	ISIA_SYNPJ
89	6	6.4	344	1	CEA6_HUMAN
90	6	6.4	346	1	DHAS_BACSU
91	6	6.4	347	1	NU2M_DICMA
92	6	6.4	349	1	SOHB_ECOLI
93	6	6.4	356	1	LPXB_AQUAE
94	6	6.4	362	1	NTCP_MOUSE
95	6	6.4	364	1	PLCE_HUMAN
96	6	6.4	365	1	FUT3_BOVIN
97	6	6.4	378	1	FETB_RAT
98	6	6.4	380	1	CYB_FONPA
99	6	6.4	380	1	CYB_PONPP
100	6	6.4	380	1	CYB_PONPY

P12838	homo sapien
P22522	escherichia
O60351	methanococc
O07591	bacillus su
Q28586	ovis aries
P52173	bos taurus
P46685	cercocobus
O02699	equus cabal
O08987	cavia porce
P71706	mycobacteri
Q9WS38	tupaia para
P11441	homo sapien
Q05557	anas platyr
Q05713	gallus gall
P41228	cricetulus
Q95971	homo sapien
Q8Y1J6	anabaena sp
O88875	mus musculus
P44088	haemophilus
P38101	escherichia
Q82hf4	yersinia pe
Q59071	methanococc
O14198	schizosacch
P50272	porphyra pu
P39541	saccharomyc
P36923	enterococcu
Q9Jtr9	neisseria m
Q9Jyt0	neisseria m
Q17025	anopheles g
Q27289	anopheles g
Q92xx8	papio hamad
P33464	porcine res
P36299	canine ente
P09175	porcine tra
P04135	porcine tra
P24412	porcine res
P25878	feline infe
P58054	escherichia
P39874	escherichia
Q95840	magnaporthe
P16701	escherichia
P79400	sus scrofa
P22190	citrobacter
Q9mxv8	sus scrofa
P71373	haemophilus
Q9per5	xyella fas
Q9nyv9	homo sapien
O67303	aquifex aeo
P59537	homo sapien
P59546	homo sapien
P33562	bacillus su
P55636	rhizobium s
P23101	pseudomonas
Q57711	methanococc
P11157	synecococc
P40199	homo sapien
Q04797	bacillus su
P41305	didelphis m
P24213	escherichia
O67420	aquifex aeo
O08705	mus musculus
Q9nuq2	homo sapien
Q1126	bos taurus
Q9qx79	rattus norv
P92701	pongo pygma
P92723	pongo pygma
Q35614	pongo pygma

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 ; Search time 44.368 Seconds
(without alignments)
546.722 Million cell updates/sec

Title: US-10-045-180A-3
Perfect score: 94
Sequence: 1 MNTLLSFLVALQWAE.....GEHLGCTCILGERYPICCY 94

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	8	8.5	128	5	Q19064
2	8	8.5	467	10	Q9FF99
3	8	8.5	649	4	Q96PL7
4	7	7.4	87	16	O24861
5	7	7.4	87	16	Q9ZN45
6	7	7.4	96	6	Q9TU00
7	7	7.4	96	6	Q9TT29
8	7	7.4	96	6	P82318
9	7	7.4	111	16	Q8EUG5
10	7	7.4	151	11	Q99KZ7
11	7	7.4	164	16	Q98MW4
12	7	7.4	196	11	Q920N0
13	7	7.4	200	2	Q9KGW7
14	7	7.4	222	12	Q9QTD0
15	7	7.4	225	10	Q8RU05
16	7	7.4	235	16	Q92TQ1

17	7	7.4	248	10	Q8LLZ3
18	7	7.4	265	16	Q92RV2
19	7	7.4	287	4	Q13984
20	7	7.4	296	16	Q8PHM1
21	7	7.4	318	16	Q8R9J7
22	7	7.4	328	10	Q9C5Z0
23	7	7.4	328	10	Q94K09
24	7	7.4	339	16	Q69943
25	7	7.4	347	16	Q83285
26	7	7.4	348	11	Q8C0P2
27	7	7.4	364	4	Q8NBA5
28	7	7.4	423	4	Q00717
29	7	7.4	430	4	Q15600
30	7	7.4	436	16	Q9XAB6
31	7	7.4	438	10	Q8LMD1
32	7	7.4	461	4	Q60430
33	7	7.4	461	4	Q13854
34	7	7.4	464	4	Q16170
35	7	7.4	468	4	Q96CA7
36	7	7.4	484	4	Q9UMN4
37	7	7.4	547	10	Q8S5D3
38	7	7.4	582	10	Q94HB7
39	7	7.4	628	16	Q9A4M9
40	7	7.4	691	10	Q943H2
41	7	7.4	702	4	Q8N4D0
42	7	7.4	712	6	Q8WMU5
43	7	7.4	714	10	Q8W514
44	7	7.4	715	10	Q8S5G0
45	7	7.4	765	10	Q8LHK4
46	7	7.4	901	5	Q9VCX7
47	7	7.4	948	5	Q95RT4
48	7	7.4	1242	5	Q9GRQ5
49	7	7.4	1481	6	Q62673
50	7	7.4	1481	6	Q9RTX9
51	7	7.4	1481	6	Q9TUQ2
52	7	7.4	1481	6	Q9TSP5
53	7	7.4	1481	6	Q62668
54	7	7.4	1485	13	P70034
55	7	7.4	1499	12	Q8U2B6
56	7	7.4	3703	5	Q8IMW2
57	7	7.4	3912	5	Q9VC56
58	6	6.4	36	8	Q959Y6
59	6	6.4	39	13	Q8UJ30
60	6	6.4	39	13	Q8UJ31
61	6	6.4	61	16	Q9RJ33
62	6	6.4	70	16	Q8Y1Z1
63	6	6.4	71	16	Q982X6
64	6	6.4	72	12	Q8JMX9
65	6	6.4	72	16	Q8G1D2
66	6	6.4	73	12	Q8QF26
67	6	6.4	77	16	Q8DAF1
68	6	6.4	82	10	Q9SDE3
69	6	6.4	85	16	Q92DH7
70	6	6.4	86	10	Q944F0
71	6	6.4	89	8	Q31693
72	6	6.4	90	5	P91520
73	6	6.4	93	16	Q8NLU4
74	6	6.4	95	2	Q44487
75	6	6.4	95	8	Q9G9T9
76	6	6.4	103	4	Q75215
77	6	6.4	106	5	O44457
78	6	6.4	107	6	Q9BG30
79	6	6.4	107	16	Q9L9J6
80	6	6.4	111	11	Q91XS4
81	6	6.4	117	16	P94484
82	6	6.4	118	6	Q9TV10
83	6	6.4	120	8	Q8SQA1
84	6	6.4	123	12	Q9QAJ3
85	6	6.4	123	12	Q9QAL4
86	6	6.4	123	12	Q9QAK8
87	6	6.4	123	12	Q9QAL2
88	6	6.4	124	2	Q45924
89	6	6.4	124	16	Q9KQ95

Q811z3	oryza sativ
Q82rv2	rhizobium m
Q13984	homo sapien
Q8phm1	xanthomonas
Q8r9j7	thermoanaer
Q9c5z0	arabidopsis
Q94k09	arabidopsis
Q69943	streptomyce
Q83285	treponema p
Q8c0p2	mus musculu
Q8nba5	homo sapien
Q00717	homo sapien
Q15600	homo sapien
Q9xaa6	streptomyce
Q8lmd1	oryza sativ
Q60430	homo sapien
Q13854	homo sapien
Q16170	homo sapien
Q96ca7	homo sapien
Q9umna	homo sapien
Q885d3	oryza sativ
Q94hb7	oryza sativ
Q9a4m9	caulobacter
Q943h2	oryza sativ
Q8n4d0	homo sapien
Q8wmu5	canis famil
Q8w514	oryza sativ
Q8s5g0	oryza sativ
Q8lhk4	oryza sativ
Q9vxt7	drosophila
Q95rt4	drosophila
Q9grq5	leishmania
Q62673	macaca mula
Q9ttx9	oryctolagus
Q9tuq2	macaca neme
Q9tsp5	papio anubi
Q62668	macaca mula
P70034	xenopus lae
Q8uzb6	grapevine f
Q8imw2	drosophila
Q9vc56	drosophila
Q959y6	gymnothorax
Q8jj30	ficedula al
Q8jj31	ficedula hy
Q9rj33	streptomyce
Q8y1z1	raletonia s
Q982x6	rhizobium l
Q8jmx9	human astro
Q8g1d2	brucella su
Q8gpz6	camelopox vi
Q8daf1	vibrio vuln
Q9sde3	oryza sativ
Q92dh7	listeria in
Q944f0	oryza sativ
Q31693	artemia par
P91520	caenorhabdi
Q8nlu4	corynebacte
Q44487	a mo-nitrog
Q9g9t9	nymphon gra
Q75215	homo sapien
O44457	caenorhabdi
Q9bg30	lemur catta
Q9l9j6	salmonella
Q91xs4	mus musculu
P94484	bacillus su
Q9tv10	canis famil
Q8sga1	anthopleura
Q9qaj3	equine arte
Q9qal4	equine arte
Q9qak8	equine arte
Q9gal2	equine arte
Q45924	clostridium
Q9kq95	vibrio chol

Q47664 escherichia
Q8gme4 sulfurospir
Q8rhd8 fusobacteri
Q8dul1 streptococc
Q8uk07 agrobacteri
Q95j76 canis famli
Q9ymy5 sus scrofa
Q42269 sus scrofa
Q69489 mycobacteri
Q8c344 mus musculu
Q9wzk4 thermotoga

ALIGNMENTS

RESULT 1

Q19064 Q19064 PRELIMINARY; PRT; 128 AA.
AC Q19064;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical 13.7 Kda protein.
GN E04F6.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Pauley A.;
RT "The sequence of C. elegans cosmid E04F6.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U28943; AAA68364.1; -.
DR WormPep; E04F6.9; CE01221.
KW Hypothetical protein.
SQ SEQUENCE 128 AA; 13694 MW; 7BBAF31DB567ACEB CRC64;

Query Match 8.5%; Score 8; DB 5; Length 128;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LSAPLLVA 14
| | | | |
Db 6 LSAPLLVA 13

RESULT 2

Q9FF99 Q9FF99 PRELIMINARY; PRT; 467 AA.
AC Q9FF99;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Amino acid transporter.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Columbia;
RX MEDLINE=97471969; PubMed=9330910;
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,
RA Miyajima N., Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence
RT features of the 1.6 Mb regions covered by twenty physically assigned
RT clones";
RL DNA Res. 4:215-230(1997).
DR EMBL; AB005244; BAB10054.1; -.
DR InterPro; IPR002422; AA/re_l_permease2.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF01490; Aa_trans_1.
DR PROSITE; PS00030; RRM_RNP_1; 1.
SQ SEQUENCE 467 AA; 51722 MW; 1BE41C85CE364406 CRC64;

Query Match 8.5%; Score 8; DB 10; Length 467;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 TLLSAFL 12
| | | | |
Db 72 TLLSAFL 79

RESULT 3

Q96PL7 Q96PL7 PRELIMINARY; PRT; 649 AA.
AC Q96PL7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Concentrative nucleoside transporter 1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RA Mata J.F., Garcia-Manteiga J.M., Lostao P., Fernandez-Veledo S.,
RA Larrayoz I.M., Lloberas J., Casado J., Pastor-Anglada M.;
RT "Role of the human concentrative nucleoside transporter (hCNT1) in the
RT cytotoxic action of 5'-deoxy-5-fluorouridine, an active intermediate
RT metabolite of capecitabine, a novel oral fluoropyrimidine used in
RT cancer treatment.";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF309632; AAL09447.1; -.
DR InterPro; IPR002668; Nucleid_transp2.
DR Pfam; PF01773; Nucleoside_tra2; 1.
DR ProDom; PD003768; Nucleid_Transp2; 1.
DR TIGRFAMs; TIGR00804; nup; 1.
SQ SEQUENCE 649 AA; 71598 MW; DCD4EAF585A00CD CRC64;

Query Match 8.5%; Score 8; DB 4; Length 649;
Best Local Similarity 100.0%; Pred. No. 9.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LSAPLLVA 14
| | | | |
Db 94 LSAPLLVA 101

RESULT 4

O24861 O24861 PRELIMINARY; PRT; 87 AA.
AC O24861;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 ; Search time 55.648 Seconds
(without alignments)
268.119 Million cell updates/sec

Title: US-10-045-180A-3

Perfect score: 94

Sequence: 1 MRTLTLISAFLLVALQWAE.....GEHLGTCFILGERYPICCY 94

Scoring table: OLIGO Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 158726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A_Geneseq_19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-embl/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	94	20	AAV01603 Amino acid sequenc
2	94	100.0	94	20	AAV01604 Amino acid sequenc
3	31	33.0	31	20	AAV01606 Fragment of human
4	8	8.5	58	22	ABB44491 Peptide #11997 enc
5	8	8.5	58	22	AAU32095 Novel human secret
6	8	8.5	58	22	AAV78244 Human bone marrow
7	8	8.5	58	23	ABG47256 Human peptide enco
8	8	8.5	59	24	AAE33811 Human mature retro
9	8	8.5	72	19	AAV71692 Human defensin pre

10	8	8.5	94	23	ABB98494	Human defensin 1.
11	8	8.5	101	23	ABB97309	Novel human protei
12	8	8.5	111	21	AAG45731	Arabidopsis thalia
13	8	8.5	118	21	AAG17608	Arabidopsis thalia
14	8	8.5	120	12	AAI15222	Chronic myelogenou
15	8	8.5	124	19	AAW71693	R21-Hpro-28 fusion
16	8	8.5	125	21	AAG45730	Arabidopsis thalia
17	8	8.5	138	23	ABP68966	Human polypeptide
18	8	8.5	144	24	AAE33860	Human retocyclin
19	8	8.5	161	23	ABP68985	Human polypeptide
20	8	8.5	168	22	AAU31085	Novel human secret
21	8	8.5	176	19	AAW71696	R78-hpro-28 fusio
22	8	8.5	196	19	AAW71694	R21-2Hpro-28 fusio
23	8	8.5	467	21	AAG45700	Arabidopsis thalia
24	8	8.5	467	23	ABP33586	Herbicidally activ
25	8	8.5	649	19	AAW70259	Human concentrati
26	8	8.5	649	19	AAW70260	Human concentrati
27	8	8.5	650	19	AAW49107	Human concentrati
28	8	8.5	702	10	AAV94014	Carcinoembryonic c
29	8	8.5	702	24	ABU04817	Human expressed pr
30	7	7.4	10	20	AAV46751	Immunogenic peptid
31	7	7.4	10	23	AAU96020	Carcino embryonic
32	7	7.4	14	22	AAH88012	CD66 peptide CD66a
33	7	7.4	14	22	AAH88105	CD66 peptide CD66e
34	7	7.4	15	21	AAV98827	HLA class II bindi
35	7	7.4	15	21	AAV98822	HLA class II bindi
36	7	7.4	43	21	AAH12435	Pro-CEMA amino aci
37	7	7.4	60	10	AAV91535	Modified region 30
38	7	7.4	65	21	AAV33208	Zea mays protein f
39	7	7.4	67	23	ABU51508	Helicobacter pylor
40	7	7.4	81	23	ABU51695	Helicobacter pylor
41	7	7.4	83	23	ABP37584	Novel human protei
42	7	7.4	87	19	AAW98778	H. pylori GHP0 118
43	7	7.4	94	14	AAH44818	Deduced sequence o
44	7	7.4	94	23	ABV79000	Transplant media a
45	7	7.4	95	23	AAU91040	Transplant media a
46	7	7.4	96	23	AAU91053	Transplant media a
47	7	7.4	96	23	AAU91054	Sequence of the ga
48	7	7.4	100	14	AAH44819	Human GENSER prote
49	7	7.4	101	24	ABP76314	Drosophila melanog
50	7	7.4	105	22	ABH61388	Human polypeptide
51	7	7.4	113	22	AAO02198	Human polypeptide
52	7	7.4	114	21	AAV74986	Neisseria meningit
53	7	7.4	115	21	AAV74984	Neisseria gonorrhoe
54	7	7.4	123	20	AAW67855	Human secreted pro
55	7	7.4	133	21	AAV74985	Neisseria meningit
56	7	7.4	170	22	AAU51905	Propionibacterium
57	7	7.4	173	13	AAH22266	Collagen type prot
58	7	7.4	196	22	AAE10311	Tamias asiaticus h
59	7	7.4	286	9	AAH81224	Carcinoembryonic a
60	7	7.4	286	15	AAH54712	Carcinoembryonic a
61	7	7.4	286	17	AAW06871	Carcinoembryonic a
62	7	7.4	286	20	AAW83136	Human CEA protein
63	7	7.4	286	24	ABU04818	Human expressed pr
64	7	7.4	286	24	ABU04824	Human expressed pr
65	7	7.4	286	24	ABU04829	Human expressed pr
66	7	7.4	287	24	ABU04801	Human expressed pr
67	7	7.4	287	24	ABU04841	Human expressed pr
68	7	7.4	321	11	AAH06434	SPI-like protein e
69	7	7.4	321	13	AAH22045	Carcino embryonic
70	7	7.4	321	24	ABU05236	Human expressed pr
71	7	7.4	323	21	AAH06220	Arabidopsis thalia
72	7	7.4	323	21	AAH07395	Arabidopsis thalia
73	7	7.4	323	21	AAH48723	Arabidopsis thalia
74	7	7.4	328	21	AAH06219	Arabidopsis thalia
75	7	7.4	328	21	AAH07394	Arabidopsis thalia
76	7	7.4	328	21	AAH48722	Arabidopsis thalia
77	7	7.4	344	9	AAH81225	Carcinoembryonic a
78	7	7.4	344	10	AAH94000	Sequence of CEA an
79	7	7.4	344	15	AAH54715	Carcinoembryonic a
80	7	7.4	344	17	AAW06874	Carcinoembryonic a
81	7	7.4	344	20	AAH83139	Human mature retro
82	7	7.4	344	24	ABU04820	Human expressed pr

83 7 7.4 344 24 ABU04826 Human expressed pr
84 7 7.4 344 24 ABU04831 Human expressed pr
85 7 7.4 349 24 ABU04803 Human expressed pr
86 7 7.4 349 24 ABU04805 Human expressed pr
87 7 7.4 349 24 ABU04807 Human expressed pr
88 7 7.4 349 24 ABU04808 Human expressed pr
89 7 7.4 349 24 ABU04835 Human expressed pr
90 7 7.4 349 24 ABU04838 Human expressed pr
91 7 7.4 351 13 AAR22318 Carcino embryonic
92 7 7.4 368 10 AAP93996 Transmembrane (TM)
93 7 7.4 369 21 AAP93996 Arabidopsis thalia
94 7 7.4 369 21 AAG06218 Arabidopsis thalia
95 7 7.4 372 24 ABU04802 Human expressed pr
96 7 7.4 417 13 AAR22044 Carcino embryonic
97 7 7.4 422 20 AAP34959 Chlamydia pneumoni
98 7 7.4 430 10 AAP93994 Sequence of transmem
99 7 7.4 451 20 AAP37158 Protein involved in
100 7 7.4 464 10 AAP93995 Transmembrane (TM)

ALIGNMENTS

RESULT 1
AAY01603
ID AAY01603 standard; protein; 94 AA.
XX AAY01603;
AC
DT 18-JUN-1999 (first entry)
XX
DE Amino acid sequence of the human defensin (Def-X) protein.
XX
KW Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
KW cytosolic; anticancer; inflammation; tissue repair;
KW endocrine regulation; corticostatic regulation; cancer; melanoma;
KW AIDS; immune deficiency; psoriasis.
XX
OS Homo sapiens.
XX
PN FR2767832-AL.
XX
PD 05-MAR-1999.
XX
PF 29-AUG-1997; 97FR-0010823.
XX
PR 29-AUG-1997; 97FR-0010823.
XX
PA (GEST) GENSET SA.
XX
PI Bougueleret L, Chumakov I;
XX
DR WPI; 1999-183266/16.
XX
PT Human defensin polypeptide Def-X - useful as antimicrobial agent,
PT anticancer agent, pesticide, etc.
XX
PS Claim 1; Fig 1; 56pp; French.
XX
CC The present sequence represents human defensin (Def-X). The Def-X
CC polypeptide can be used as an antimicrobial, antiparasitic agent or
CC a pesticide. The Def-X polypeptide can be used as a cytostatic
CC (especially anticancer) agent, and as an agent for modulating processes
CC of inflammation, tissue repair and endocrine (especially corticostatic)
CC regulation. The polypeptide can be used in a composition for external
CC topical use, especially in a cosmetic composition. Compositions
CC containing Def-X can be used for prevention and treatment of microbial
CC and parasitic infections, especially where the microbial or parasitic
CC infections are Gram-positive or -negative bacterial infections or
CC mycobacterial, fungal or spirochaete infections, or where the viral
CC infections are associated with enveloped viruses, especially HSV and HIV.
CC The compositions can be used for prevention and/or treatment of cancers,
CC especially melanomas, or liver cancer, prostate cancer, non-small-cell
CC lung cancer or colorectal carcinoma, and for enhancing immunity,

CC lung cancer or colorectal carcinoma, and for enhancing immunity,
CC especially in the case of AIDS, or preventing immune deficiency,
CC especially in the treatment of psoriasis, as well as for modulating
CC inflammatory processes, especially in the case of chronic inflammatory
CC disorders.
XX
SQ Sequence 94 AA;
Query Match 100.0%; Score 94; DB 20; Length 94;
Best Local Similarity 100.0%; Pred. NO. 2.5e-91;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRTLTLTSLAFLVALQAWAEPLQARAHENPAOKPPADDDVVIYPSGDDSCSLQVPGST 60
Db 1 MRTLTLTSLAFLVALQAWAEPLQARAHENPAOKPPADDDVVIYPSGDDSCSLQVPGST 60
QY 61 KGLICHRCVLYCIFGHEHLGTCFILGERYPICCY 94
Db 61 KGLICHRCVLYCIFGHEHLGTCFILGERYPICCY 94
RESULT 2
AAY01604
ID AAY01604 standard; protein; 94 AA.
XX
AC AAY01604;
XX
DT 18-JUN-1999 (first entry)
XX
DE Amino acid sequence of the human defensin (Def-X) protein.
XX
KW Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
KW cytosolic; anticancer; inflammation; tissue repair;
KW endocrine regulation; corticostatic regulation; cancer; melanoma;
KW AIDS; immune deficiency; psoriasis.
XX
OS Homo sapiens.
XX
PN FR2767832-AL.
XX
PD 05-MAR-1999.
XX
PF 29-AUG-1997; 97FR-0010823.
XX
PR 29-AUG-1997; 97FR-0010823.
XX
PA (GEST) GENSET SA.
XX
PI Bougueleret L, Chumakov I;
XX
DR WPI; 1999-183266/16.
XX
PT N-PSDB; AAX26697.
XX
PT Human defensin polypeptide Def-X - useful as antimicrobial agent,
PT anticancer agent, pesticide, etc.
XX
PS Disclosure; Fig 3; 56pp; French.
XX
CC The present sequence represents human defensin (Def-X). The Def-X
CC polypeptide can be used as an antimicrobial, antiparasitic agent or
CC a pesticide. The Def-X polypeptide can be used as a cytostatic
CC (especially anticancer) agent, and as an agent for modulating processes
CC of inflammation, tissue repair and endocrine (especially corticostatic)
CC regulation. The polypeptide can be used in a composition for external
CC topical use, especially in a cosmetic composition. Compositions
CC containing Def-X can be used for prevention and treatment of microbial
CC and parasitic infections, especially where the microbial or parasitic
CC infections are Gram-positive or -negative bacterial infections or
CC mycobacterial, fungal or spirochaete infections, or where the viral
CC infections are associated with enveloped viruses, especially HSV and HIV.
CC The compositions can be used for prevention and/or treatment of cancers,
CC especially melanomas, or liver cancer, prostate cancer, non-small-cell
CC lung cancer or colorectal carcinoma, and for enhancing immunity,

CC especially in the case of AIDS, or preventing immune deficiency.
 CC especially in the treatment of psoriasis, as well as for modulating
 CC inflammatory processes, especially in the case of chronic inflammatory
 CC disorders.

XX
 SQ Sequence 94 AA;
 Query Match 100.0%; Score 94; DB 20; Length 94;
 Best Local Similarity 100.0%; Pred. No. 2.5e-91;
 Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRTLISAFLLVALQAWAEPLQARAHMPAQKPPADDQDVVYFSGDDSCSLQVPGST 60
 |||||
 DB 1 MRTLISAFLLVALQAWAEPLQARAHMPAQKPPADDQDVVYFSGDDSCSLQVPGST 60
 |||||

QY 61 KGLICHRCVLYCIFGEHLGGTCFILGERYPICCY 94
 |||||
 DB 61 KGLICHRCVLYCIFGEHLGGTCFILGERYPICCY 94
 |||||

RESULT 3
 ID AAY01606 standard; peptide; 31 AA.
 XX
 AC AAY01606;
 XX
 DT 18-JUN-1999 (first entry)
 XX
 DE Fragment of human defensin (Def-X) protein.
 XX
 KW Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
 KW cytostatic; anticancer; inflammation; tissue repair;
 KW endocrine regulation; corticostatic regulation; cancer; melanoma;
 KW AIDS; immune deficiency; psoriasis.
 OS Homo sapiens.
 XX
 XX FR2767832-Al.
 XX
 PD 05-MAR-1999.
 XX
 XX 29-AUG-1997; 97FR-0010823.
 PF
 XX 29-AUG-1997; 97FR-0010823.
 XX
 PA (GEST) GENSET SA.
 XX
 PI Bougueleret L, Chumakov I;
 XX
 DR WPI; 1999-183266/16.
 XX
 PT Human defensin polypeptide Def-X - useful as antimicrobial agent,
 PT anticancer agent, pesticide, etc.
 XX
 PS Claim 2; Page 33; 56pp; French.
 XX

The present sequence represents human defensin (Def-X) fragment. The Def-X polypeptide can be used as an antimicrobial, antiparasitic agent or a pesticide. The Def-X polypeptide can be used as a cytostatic (especially anticancer) agent, and as an agent for modulating processes of inflammation, tissue repair and endocrine (especially corticostatic) regulation. The polypeptide can be used in a composition for external topical use, especially in a cosmetic composition. Compositions containing Def-X can be used for prevention and treatment of microbial and parasitic infections, especially where the microbial or parasitic infections are Gram-positive or -negative bacterial infections or mycobacterial, fungal or spirochaete infections, or where the viral infections are associated with enveloped viruses, especially HSV and HIV. The compositions can be used for prevention and/or treatment of cancers, especially melanomas, or liver cancer, prostate cancer, non-small-cell lung cancer or colorectal carcinoma, and for enhancing immunity, especially in the case of AIDS, or preventing immune deficiency, especially in the treatment of psoriasis, as well as for modulating

CC inflammatory processes, especially in the case of chronic inflammatory
 CC disorders.

XX
 SQ Sequence 31 AA;
 Query Match 33.0%; Score 31; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 3.2e-25;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 ICHCRVLYCIFGEHLGGTCFILGERYPICCY 94
 |||||
 DB 1 ICHCRVLYCIFGEHLGGTCFILGERYPICCY 31
 |||||

RESULT 4
 ID ABB44491 standard; Peptide; 58 AA.
 XX
 AC ABB44491;
 XX
 DT 04-FEB-2002 (first entry)
 XX
 DE Peptide #11997 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human foetal liver -
 XX
 PS Claim 27; SEQ ID NO 37126; 639pp + sequence listing; English.
 XX

The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a peptide encoded by a single exon nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 58 AA;
 Query Match 8.5%; Score 8; DB 22; Length 58;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 AEPLQARA 26
 |||||
 DB 19 AEPLQARA 26
 |||||

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:51:01 ; Search time 37.6 Seconds
(without alignments)
466.897 Million cell updates/sec

Title: US-10-045-180A-3

Perfect score: 94

Sequence: I MRTLTLISAFLLVALQWAE.....GEHLGGTCFLGERYPICCY 94

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 696363 seqs, 186758610 residues

Word size : 0

Total number of hits satisfying chosen parameters: 696363

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Published Applications AA.*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/FCI_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/FCI05_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	94	100.0	94	14	US-10-013-770-3
2	94	100.0	94	14	US-10-045-180A-3
3	44	46.8	44	14	US-10-013-770-5
4	44	46.8	44	14	US-10-045-180A-5
5	31	33.0	31	14	US-10-013-770-6
6	31	33.0	31	14	US-10-045-180A-6
7	19	20.2	19	14	US-10-013-770-4
8	19	20.2	19	14	US-10-045-180A-4
9	8	8.5	58	9	US-09-864-761-47890
10	8	8.5	59	12	US-10-141-645-12
11	8	8.5	94	14	US-10-045-180A-12
12	8	8.5	118	12	US-10-141-645-11
13	8	8.5	140	12	US-10-141-645-65
14	8	8.5	140	12	US-10-141-645-66
15	8	8.5	140	12	US-10-141-645-69
16	8	8.5	140	12	US-10-141-645-69
17	8	8.5	140	12	US-10-141-645-69
18	8	8.5	140	12	US-10-141-645-69
19	8	8.5	140	12	US-10-141-645-69
20	8	8.5	140	12	US-10-141-645-69
21	8	8.5	140	12	US-10-141-645-69
22	8	8.5	140	12	US-10-141-645-69
23	8	8.5	140	12	US-10-141-645-69
24	8	8.5	140	12	US-10-141-645-69
25	8	8.5	140	12	US-10-141-645-69
26	8	8.5	140	12	US-10-141-645-69
27	8	8.5	140	12	US-10-141-645-69
28	8	8.5	140	12	US-10-141-645-69
29	8	8.5	140	12	US-10-141-645-69
30	8	8.5	140	12	US-10-141-645-69
31	8	8.5	140	12	US-10-141-645-69
32	8	8.5	140	12	US-10-141-645-69
33	8	8.5	140	12	US-10-141-645-69
34	8	8.5	140	12	US-10-141-645-69
35	8	8.5	140	12	US-10-141-645-69
36	8	8.5	140	12	US-10-141-645-69
37	8	8.5	140	12	US-10-141-645-69
38	8	8.5	140	12	US-10-141-645-69
39	8	8.5	140	12	US-10-141-645-69
40	8	8.5	140	12	US-10-141-645-69
41	8	8.5	140	12	US-10-141-645-69
42	8	8.5	140	12	US-10-141-645-69
43	8	8.5	140	12	US-10-141-645-69
44	8	8.5	140	12	US-10-141-645-69
45	8	8.5	140	12	US-10-141-645-69
46	8	8.5	140	12	US-10-141-645-69
47	8	8.5	140	12	US-10-141-645-69
48	8	8.5	140	12	US-10-141-645-69
49	8	8.5	140	12	US-10-141-645-69
50	8	8.5	140	12	US-10-141-645-69
51	8	8.5	140	12	US-10-141-645-69
52	8	8.5	140	12	US-10-141-645-69
53	8	8.5	140	12	US-10-141-645-69
54	8	8.5	140	12	US-10-141-645-69
55	8	8.5	140	12	US-10-141-645-69
56	8	8.5	140	12	US-10-141-645-69
57	8	8.5	140	12	US-10-141-645-69
58	8	8.5	140	12	US-10-141-645-69
59	8	8.5	140	12	US-10-141-645-69
60	8	8.5	140	12	US-10-141-645-69
61	8	8.5	140	12	US-10-141-645-69
62	8	8.5	140	12	US-10-141-645-69
63	8	8.5	140	12	US-10-141-645-69
64	8	8.5	140	12	US-10-141-645-69
65	8	8.5	140	12	US-10-141-645-69
66	8	8.5	140	12	US-10-141-645-69
67	8	8.5	140	12	US-10-141-645-69
68	8	8.5	140	12	US-10-141-645-69
69	8	8.5	140	12	US-10-141-645-69
70	8	8.5	140	12	US-10-141-645-69
71	8	8.5	140	12	US-10-141-645-69
72	8	8.5	140	12	US-10-141-645-69
73	8	8.5	140	12	US-10-141-645-69
74	8	8.5	140	12	US-10-141-645-69
75	8	8.5	140	12	US-10-141-645-69
76	8	8.5	140	12	US-10-141-645-69
77	8	8.5	140	12	US-10-141-645-69
78	8	8.5	140	12	US-10-141-645-69
79	8	8.5	140	12	US-10-141-645-69
80	8	8.5	140	12	US-10-141-645-69
81	8	8.5	140	12	US-10-141-645-69
82	8	8.5	140	12	US-10-141-645-69
83	8	8.5	140	12	US-10-141-645-69
84	8	8.5	140	12	US-10-141-645-69
85	8	8.5	140	12	US-10-141-645-69
86	8	8.5	140	12	US-10-141-645-69
87	8	8.5	140	12	US-10-141-645-69
88	8	8.5	140	12	US-10-141-645-69
89	8	8.5	140	12	US-10-141-645-69
90	8	8.5	140	12	US-10-141-645-69
91	8	8.5	140	12	US-10-141-645-69
92	8	8.5	140	12	US-10-141-645-69
93	8	8.5	140	12	US-10-141-645-69
94	8	8.5	140	12	US-10-141-645-69
95	8	8.5	140	12	US-10-141-645-69
96	8	8.5	140	12	US-10-141-645-69
97	8	8.5	140	12	US-10-141-645-69
98	8	8.5	140	12	US-10-141-645-69
99	8	8.5	140	12	US-10-141-645-69
100	8	8.5	140	12	US-10-141-645-69

Sequence 70, Appl
Sequence 71, Appl
Sequence 123, Appl
Sequence 125, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 28, Appl
Sequence 31739, A
Sequence 28, Appl
Sequence 10, Appl
Sequence 76, Appl
Sequence 89, Appl
Sequence 90, Appl
Sequence 11, Appl
Sequence 169, Appl
Sequence 169, Appl
Sequence 169, Appl
Sequence 169, Appl
Sequence 169, Appl
Sequence 3, Appl
Sequence 14932, A
Sequence 4871, Ap
Sequence 592, App
Sequence 341, App
Sequence 87, Appl
Sequence 17, Appl
Sequence 1133, Ap
Sequence 4, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 6, Appl
Sequence 2, Appl
Sequence 72, Appl
Sequence 14, Appl
Sequence 13, Appl
Sequence 2879, Ap
Sequence 13958, A
Sequence 131, App
Sequence 131, App
Sequence 130, App
Sequence 130, App
Sequence 125, App
Sequence 125, App
Sequence 146, App
Sequence 446, App
Sequence 449, App
Sequence 12, Appl
Sequence 8, Appl
Sequence 129, App
Sequence 129, App
Sequence 129, App
Sequence 370, App
Sequence 6630, Ap
Sequence 132, App
Sequence 132, App
Sequence 13, Appl
Sequence 23, Appl
Sequence 9, Appl
Sequence 25, Appl
Sequence 25, Appl
Sequence 43600, A
Sequence 1191, Ap
Sequence 39, Appl
Sequence 1191, Ap
Sequence 28, Appl
Sequence 29, Appl
Sequence 16, Appl
Sequence 7, Appl
Sequence 10, Appl
Sequence 10, Appl
Sequence 86, Appl
Sequence 86, Appl
Sequence 403, App
Sequence 79, Appl
Sequence 5, Appl

89 Sequence 81, Appl
90 Sequence 81, Appl
91 Sequence 7570, Ap
92 Sequence 351, App
93 Sequence 4179, Ap
94 Sequence 82, Appl
95 Sequence 163, App
96 Sequence 37898, A
97 Sequence 2, Appli
98 Sequence 133, App
99 Sequence 133, App
100 Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-10-013-770-3
; Sequence 3, Application US/10013770
; Publication No. US20020115151A1
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; THERAPEUTIC APPLICATIONS

NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 550 West C Street
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/013,770
FILING DATE: 10-Dec-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/486,580
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Hart, Daniel
REGISTRATION NUMBER: 40,637
REFERENCE/DOCKET NUMBER: GENSET.064C1
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 AMINO ACIDS
TYPE: AMINO ACID
STRANDEDNESS: SINGLE
TOPOLOGY: LINEAR
MOLECULE TYPE: PROTEIN
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIGNAL PEPTIDE
LOCATION: 1..19
FEATURE:
NAME/KEY: PRO REGION
LOCATION: 20..63
FEATURE:
NAME/KEY: MATURE PEPTIDE
LOCATION: 64..94
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-10-013-770-3
Query Match 100.0%; Score 94; DB 14; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-88;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRTLTLISAFLLVALQAWAEPLQARAHMPAKQPPADDDQVVIYFSGDDSCSLQVPGST 60
DB 1 MRTLTLISAFLLVALQAWAEPLQARAHMPAKQPPADDDQVVIYFSGDDSCSLQVPGST 60
QY 61 KGLICHRCRVLYCIFGEHLGGTCTFILGERYPICCY 94
DB 61 KGLICHRCRVLYCIFGEHLGGTCTFILGERYPICCY 94

RESULT 2
US-10-045-180A-3
; Sequence 3, Application US/10045180A
; Publication No. US20020182703A1
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Compositi
; FILE REFERENCE: GEN-100D1
; CURRENT FILING DATE: 2001-10-18
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: US 09/486,580
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: PCT/FR98/01864
; PRIOR FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (1)..(94)
OTHER INFORMATION: Def-X preproprotein sequence
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)..(19)
OTHER INFORMATION: Def-X signal peptide
FEATURE:
NAME/KEY: PROPEP
LOCATION: (20)..(63)
OTHER INFORMATION: Def-X propeptide
FEATURE:
NAME/KEY: PEPTIDE
LOCATION: (64)..(94)
OTHER INFORMATION: Def-X mature peptide
US-10-045-180A-3

Query Match 100.0%; Score 94; DB 14; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.8e-88;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRTLTLISAFLLVALQAWAEPLQARAHMPAKQPPADDDQVVIYFSGDDSCSLQVPGST 60
DB 1 MRTLTLISAFLLVALQAWAEPLQARAHMPAKQPPADDDQVVIYFSGDDSCSLQVPGST 60

QY 61 KGLICHRCRVLYCIFGEHLGGTCTFILGERYPICCY 94
DB 61 KGLICHRCRVLYCIFGEHLGGTCTFILGERYPICCY 94

RESULT 3
US-10-013-770-5
; Sequence 5, Application US/10013770
; Publication No. US20020115151A1
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 ; Search time 21.056 Seconds
(without alignments)
188.888 Million cell updates/sec

Title: US-10-045-180A-3

Perfect score: 94

Sequence: 1 MRTLTLISAFLLVALQAWAE.....GEHLGGTCFLGGRYPICCY 94

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA:*
1: /cgm2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgm2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgm2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgm2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgm2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgm2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	94	100.0	94	4	US-09-486-580A-3
2	44	46.8	44	4	US-03-486-580A-5
3	31	33.0	31	4	US-03-486-580A-6
4	19	20.2	19	4	US-09-486-580A-4
5	8	8.5	72	3	US-09-042-071-38
6	8	8.5	94	1	US-08-158-189-12
7	8	8.5	94	1	US-08-158-189-13
8	8	8.5	124	3	US-09-042-071-39
9	8	8.5	178	3	US-03-042-071-42
10	8	8.5	195	3	US-09-042-071-40
11	8	8.5	649	3	US-08-800-291B-5
12	8	8.5	649	3	US-08-800-291B-6
13	8	8.5	650	3	US-08-800-291B-4
14	7	7.4	19	4	US-09-351-657A-42
15	7	7.4	19	4	US-09-351-657A-45
16	7	7.4	45	4	US-09-351-657A-46
17	7	7.4	45	4	US-08-158-189-2
18	7	7.4	58	1	US-08-158-189-5
19	7	7.4	95	1	US-08-158-189-14
20	7	7.4	95	1	US-08-158-189-15
21	7	7.4	100	1	US-08-158-189-7
22	7	7.4	100	1	US-08-158-189-17
23	7	7.4	111	1	US-08-466-886-19
24	7	7.4	111	3	US-08-469-617-19
25	7	7.4	284	4	US-09-252-991A-32586
26	7	7.4	309	1	US-08-463-052B-8
27	7	7.4	309	2	US-08-460-907B-8

28	7	7.4	321	6	5169835-17	Patent No. 5169835
29	7	7.4	373	4	US-09-252-991A-30414	Sequence 30414, A
30	7	7.4	422	4	US-09-198-452A-377	Sequence 377, App
31	7	7.4	464	2	US-08-602-725-32	Sequence 32, Appl
32	7	7.4	642	1	US-08-217-299-1	Sequence 1, Appl
33	7	7.4	698	2	US-08-602-725-36	Sequence 36, Appl
34	7	7.4	734	2	US-08-389-459A-17	Sequence 17, Appl
35	7	7.4	734	3	US-08-987-867A-17	Sequence 17, Appl
36	7	7.4	1476	3	US-09-256-703-2	Sequence 2, Appl
37	7	7.4	1479	2	US-08-951-912-4	Sequence 4, Appl
38	7	7.4	1479	4	US-09-174-077-4	Sequence 4, Appl
39	7	7.4	1480	1	US-07-637-621-2	Sequence 2, Appl
40	7	7.4	1480	1	US-08-136-742A-2	Sequence 2, Appl
41	7	7.4	1480	1	US-08-135-809A-2	Sequence 2, Appl
42	7	7.4	1480	1	US-08-466-886-17	Sequence 17, Appl
43	7	7.4	1480	2	US-08-951-912-2	Sequence 2, Appl
44	7	7.4	1480	2	US-08-951-912-6	Sequence 6, Appl
45	7	7.4	1480	2	US-08-469-461-2	Sequence 2, Appl
46	7	7.4	1480	2	US-08-469-461-4	Sequence 4, Appl
47	7	7.4	1480	2	US-08-691-605-2	Sequence 2, Appl
48	7	7.4	1480	2	US-08-455-552A-14	Sequence 14, Appl
49	7	7.4	1480	3	US-07-890-609-2	Sequence 2, Appl
50	7	7.4	1480	3	US-07-890-609-4	Sequence 4, Appl
51	7	7.4	1480	3	US-09-248-026-2	Sequence 2, Appl
52	7	7.4	1480	3	US-08-469-617-17	Sequence 17, Appl
53	7	7.4	1480	3	US-08-681-838A-2	Sequence 2, Appl
54	7	7.4	1480	3	US-08-681-838A-3	Sequence 3, Appl
55	7	7.4	1480	4	US-09-174-077-2	Sequence 2, Appl
56	7	7.4	1480	4	US-09-174-077-6	Sequence 6, Appl
57	7	7.4	1480	4	US-09-425-453A-2	Sequence 2, Appl
58	7	7.4	1480	4	US-09-425-453A-4	Sequence 4, Appl
59	7	7.4	1480	4	US-09-425-453A-6	Sequence 6, Appl
60	7	7.4	1480	4	US-09-425-453A-8	Sequence 8, Appl
61	7	7.4	1480	4	US-09-425-453A-10	Sequence 10, Appl
62	7	7.4	1480	4	US-09-425-453A-12	Sequence 12, Appl
63	7	7.4	1480	4	US-09-425-453A-14	Sequence 14, Appl
64	7	7.4	1480	4	US-09-425-453A-16	Sequence 16, Appl
65	7	7.4	1480	4	US-09-425-453A-18	Sequence 18, Appl
66	7	7.4	1480	4	US-09-425-453A-20	Sequence 20, Appl
67	7	7.4	1480	5	PCT-US93-11667-2	Sequence 2, Appl
68	7	7.4	1480	6	5240846-5	Patent No. 5240846
69	6	6.4	15	4	US-08-924-629C-14	Sequence 14, Appl
70	6	6.4	21	4	US-08-924-629C-13	Sequence 13, Appl
71	6	6.4	44	1	US-08-479-577-2	Sequence 2, Appl
72	6	6.4	44	2	US-08-756-977-2	Sequence 2, Appl
73	6	6.4	44	5	PCT-US93-05331-6	Sequence 6, Appl
74	6	6.4	58	4	US-09-082-358B-82	Sequence 82, Appl
75	6	6.4	66	3	US-09-338-907-131	Sequence 131, App
76	6	6.4	66	4	US-09-218-207-131	Sequence 131, App
77	6	6.4	68	3	US-09-338-907-130	Sequence 130, App
78	6	6.4	68	4	US-09-218-207-130	Sequence 130, App
79	6	6.4	77	3	US-09-338-907-125	Sequence 125, App
80	6	6.4	77	4	US-09-218-207-125	Sequence 125, App
81	6	6.4	81	4	US-09-134-001C-3127	Sequence 3127, Ap
82	6	6.4	82	4	US-09-249-542-8	Sequence 8, Appl
83	6	6.4	90	3	US-09-338-907-129	Sequence 129, App
84	6	6.4	90	4	US-09-218-207-129	Sequence 129, App
85	6	6.4	97	3	US-09-338-907-132	Sequence 132, App
86	6	6.4	97	4	US-09-309-487-23	Sequence 23, Appl
87	6	6.4	97	4	US-09-218-207-132	Sequence 132, App
88	6	6.4	97	4	US-09-967-808-23	Sequence 23, Appl
89	6	6.4	103	4	US-08-924-629C-39	Sequence 39, Appl
90	6	6.4	111	4	US-09-249-542-16	Sequence 16, Appl
91	6	6.4	111	4	US-09-107-532A-5707	Sequence 5707, Ap
92	6	6.4	112	4	US-09-249-542-7	Sequence 7, Appl
93	6	6.4	112	4	US-09-249-542-10	Sequence 10, Appl
94	6	6.4	115	4	US-09-322-409-86	Sequence 86, Appl
95	6	6.4	115	4	US-09-451-527-86	Sequence 86, Appl
96	6	6.4	125	4	US-09-199-637A-403	Sequence 403, App
97	6	6.4	128	4	US-09-199-637A-79	Sequence 79, Appl
98	6	6.4	134	4	US-09-322-409-81	Sequence 81, Appl
99	6	6.4	134	4	US-09-451-527-81	Sequence 81, Appl
100	6	6.4	134	4	US-09-371-615A-2	Sequence 2, Appl

ALIGNMENTS

```
RESULT 1
US-09-486-580A-3
; Sequence 3, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 94 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL PEPTIDE
; LOCATION: 1..19
; FEATURE:
; NAME/KEY: PRO REGION
; LOCATION: 20..63
; FEATURE:
; NAME/KEY: MATURE PEPTIDE
; LOCATION: 64..94
US-09-486-580A-3
Query Match 100.0%; Score 94; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 5e-92;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRTLTLISAFLLVALQAWAEPLQARAHMPAQKPPADDQDVVIYFSGDDSCSLQVPGST 60
Db 1 MRTLTLISAFLLVALQAWAEPLQARAHMPAQKPPADDQDVVIYFSGDDSCSLQVPGST 60

QY 61 KGLICHRVLYCIFGHELGCTCFILGERYPICCY 94
Db 61 KGLICHRVLYCIFGHELGCTCFILGERYPICCY 94

RESULT 2
US-09-486-580A-5
; Sequence 5, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
```

```
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PRO REGION
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-486-580A-5
Query Match 46.8%; Score 44; DB 4; Length 44;
Best Local Similarity 100.0%; Pred. No. 1.9e-39;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 20 EPLQARAHMPAQKPPADDQDVVIYFSGDDSCSLQVFGSTKGL 63
Db 1 EPLQARAHMPAQKPPADDQDVVIYFSGDDSCSLQVFGSTKGL 44

RESULT 3
US-09-486-580A-6
; Sequence 6, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Knobbe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 44 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PRO REGION
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL PEPTIDE
; LOCATION: 1..19
; FEATURE:
; NAME/KEY: PRO REGION
; LOCATION: 20..63
; FEATURE:
; NAME/KEY: MATURE PEPTIDE
; LOCATION: 64..94
US-09-486-580A-6
Query Match 100.0%; Score 94; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 5e-92;
Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRTLTLISAFLLVALQAWAEPLQARAHMPAQKPPADDQDVVIYFSGDDSCSLQVPGST 60
Db 1 MRTLTLISAFLLVALQAWAEPLQARAHMPAQKPPADDQDVVIYFSGDDSCSLQVPGST 60

QY 61 KGLICHRVLYCIFGHELGCTCFILGERYPICCY 94
Db 61 KGLICHRVLYCIFGHELGCTCFILGERYPICCY 94

RESULT 2
US-09-486-580A-5
; Sequence 5, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 ; Search time 6.696 Seconds
(without alignments)
445.225 Million cell updates/sec

Title: US-10-045-180A-6
Perfect score: 31
Sequence: 1 ICHCRVLYCIFGEHLGGTCFIIIGERYPICCY 31

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283308 seqs, 96168682 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

PIR 76: *
1: Pir1: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	6	19.4	115	1	G64304	hypothetical prote
2	6	19.4	124	2	H82117	hypothetical prote
3	6	19.4	131	2	AG3198	hypothetical prote
4	6	19.4	137	2	A70913	hypothetical prote
5	6	19.4	174	2	A49181	alpha B-crystallin
6	6	19.4	174	2	S58758	alpha-crystallin c
7	6	19.4	212	2	AG0854	probable sugar ald
8	6	19.4	268	2	A97260	dihydropterate sy
9	6	19.4	287	2	C90023	hypothetical prote
10	6	19.4	294	2	S19006	beta-lactamase (EC
11	6	19.4	296	2	A64110	cell division inhi
12	6	19.4	300	2	AF1286	CDP-abequose synth
13	6	19.4	300	2	A11657	CDP-abequose synth
14	6	19.4	352	2	E87338	asparaginyl-tRNA s
15	6	19.4	356	2	B70424	lipid A disacchari
16	6	19.4	386	2	B72060	2-component regula
17	6	19.4	386	2	G86563	2-component regula
18	6	19.4	408	2	A70594	probable manA prot
19	6	19.4	410	2	F87004	probable mannose-6
20	6	19.4	456	2	A64052	glutathione-disulf
21	6	19.4	491	2	D83353	mannitol dehydroge
22	6	19.4	499	2	AH2761	penicillin-binding
23	6	19.4	506	2	G97542	penicillin-binding
24	6	19.4	510	2	A55207	glycerol-3-phospha
25	6	19.4	512	2	B83197	glycerol-3-phospha
26	6	19.4	518	2	B69981	N-acetylmuramoyl-L
27	6	19.4	543	2	T39345	probable metabolit
28	6	19.4	543	2	B53978	protein-tyrosine-p
29	6	19.4	639	2	T46577	arylsulfatase (EC

protein-tyrosine-p
probable glycosyl
translation elonga
DNA uptake protein
hypothetical prote
protein F41H10.4 [
hypothetical prote
hypothetical prote
pepJ protein - Aci
neurotoxin Os-3 -
glutathione-disulf
hypothetical prote
hypothetical prote
hypothetical prote
neuropeptide Y pre
conserved hypothet
prostatic steroid-
nonhistone chromos
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
probable membrane
viJp protein - Esc
hypothetical prote
hypothetical prote
probable membrane
hypothetical prote
hypothetical prote
cytochrome c' - Rh
cytochrome c' - Rh
cytochrome c' - Rh
conserved hypothet
cellulase (EC 3.2.
profilin - Para ru
conserved hypothet
histone H3.4 - mou
hypothetical prote
hypothetical prote
hypothetical prote
hypothetical prote
two component resp
hypothetical prote
actinorhodin polyk
hypothetical cytos
conserved hypothet
hypothetical prote
conserved hypothet
hypothetical prote
hypothetical prote
hypothetical prote
disease resistance
calcium channel be
hypothetical prote
alpha-crystallin B
hypothetical prote
ribosomal protein
NADH dehydrogenase
NADH dehydrogenase
NADH dehydrogenase
NADH dehydrogenase
hypothetical prote
conserved hypothet
alpha-crystallin c
alpha-crystallin c
alpha-crystallin c
alpha-crystallin c
alpha B-crystallin

ALIGNMENTS

RESULT 1
G64304
hypothetical protein MJ0039 - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Mar-2003
C:Accession: G64304
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, R.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: G64304
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-115 <BUL>
A:Cross-references: GB:U67462; GB:L77117; NID:gi590835; PIDN:AAB98020.1; PID:gl590840; T
C:Genetics:
A:Map position: REV39998-39651
A:Start codon: TTG
C:Superfamily: RNA polymerase, subunit F

Query Match 19.4%; Score 6; DB 1; Length 115;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 ILGERY 26
Db 14 ILGERY 19
RESULT 2
H82117
hypothetical protein VC2105 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C:Accession: H82117
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, F l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: H82035; MUID:20406833; PMID:10952301
A:Accession: H82117
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-124 <HEI>
A:Cross-references: GB:AE004284; GB:AE003852; NID:g9656649; PIDN:AAF95251.1; GSPDB:GN001
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC2105
A:Map position: 1

Query Match 19.4%; Score 6; DB 2; Length 124;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 21 ILGERY 26
Db 78 ILGERY 83

RESULT 3
AG3198
hypothetical protein Atus317 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C:Accession: AG3198
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AG3198
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-131 <KUR>
A:Cross-references: GB:AE008687; PIDN:AAL46005.1; PID:gi7743760; GSPDB:GN00188
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atus317
A:Genome: plasmid

Query Match 19.4%; Score 6; DB 2; Length 131;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 12 GEHLGG 17
Db 52 GEHLGG 57

RESULT 4
A70913
hypothetical protein Rv0049 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: A70913
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S Rajadream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sqaes, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the Biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: A70913
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-137 <COL>
A:Cross-references: GB:280775; GB:AL123456; NID:g3250715; PIDN:CAB02528.1; PID:gl568588
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv0049
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0049

Query Match 19.4%; Score 6; DB 2; Length 137;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 11 FGEHLG 16
Db 61 FGEHLG 66

RESULT 5
A49181
alpha B-crystallin - chicken
C:Species: Gallus gallus (chicken)
C>Date: 19-Dec-1993 #sequence_revision 19-Nov-1994 #text_change 13-Aug-1999
C:Accession: A49181
R:Sawada, K.; Agata, K.; Eguchi, G.
Exp. Eye Res. 55, 879-887, 1992
A:Title: Crystallin gene expression in the process of lentoidogenesis in cultures of ch
A:Reference number: A49181; MUID:93137981; PMID:1283129
A:Accession: A49181
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-174 <SAW>
A:Cross-references: GB:S53164; NID:g263999; PIDN:AAB25041.1; PID:g264000

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 ; Search time 5.208 Seconds
(without alignments)
279.921 Million cell updates/sec

Title: US-10-045-180A-6

Perfect score: 31

Sequence: 1 ICHCRVLXCIHGLHGTCTILGERYPICCY 31

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 127863 seqs, 47026705 residues

Word size : 0

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	19.4	52	1 CRAB_COLLI	O12987 columba liv
2	6	19.4	52	1 CRAB_TURME	O12995 turdus meru
3	6	19.4	115	1 Y039_METUA	Q60351 methanococc
4	6	19.4	137	1 Y049_MYTCT	P71706 mycobacteri
5	6	19.4	174	1 CRAB_ANAPL	Q05557 anas platyr
6	6	19.4	174	1 CRAB_CHICK	Q05713 gallus gall
7	6	19.4	294	1 BLAC_CITDI	P22390 citrobacter
8	6	19.4	296	1 YFCH_HAEIN	P71373 haemophilus
9	6	19.4	356	1 LPXB_AQUAE	O67420 aquifex aeo
10	6	19.4	456	1 GSHR_HAEIN	P43783 haemophilus
11	6	19.4	512	1 GLPD_PSEAE	P52111 pseudomonas
12	6	19.4	813	1 EFG2_SCHPO	O94429 schizosacch
13	6	19.4	4655	1 LRP2_HUMAN	P81164 homo sapien
14	5	16.1	28	1 PHVB_ASPFI	P81440 aspergillus
15	5	16.1	35	1 Y847_BORBU	O51787 borrelia bu
16	5	16.1	52	1 CRAB_EUDEL	O12991 eudromia el
17	5	16.1	52	1 CRAB_TRASC	Q91518 trachenys s
18	5	16.1	67	1 SCX3_ORTSC	P15225 orthochirus
19	5	16.1	84	1 Y074_RHIME	P42878 rhizobium m
20	5	16.1	93	1 VE7_HPVS0	Q80928 human papil
21	5	16.1	95	1 PSC3_RAT	P02780 rattus norv
22	5	16.1	96	1 CUB5_HUMAN	P59088 homo sapien
23	5	16.1	99	1 NHPB_YEAST	P11633 saccharomyc
24	5	16.1	110	1 US36_HCWVA	P16842 human cytom
25	5	16.1	119	1 YGH2_YEAST	P53161 saccharomyc
26	5	16.1	119	1 YIJD_ECOLI	P27308 escherichia
27	5	16.1	121	1 YIPP_DROME	Q9xx40 drosophila
28	5	16.1	125	1 CYCP_RHOPH	P00145 rhodospiril
29	5	16.1	126	1 CYCP_RHORU	P00146 rhodospaudo
30	5	16.1	128	1 CYCP_RHOSP	P00147 rhodobacter
31	5	16.1	129	1 CYCP_RHOCA	P23666 persea amer
32	5	16.1	130	1 GUN2_PERA	O65812 hevea bras
33	5	16.1	131	1 PRO1_HEVBR	

ALIGNMENTS

RESULT 1

34	5	16.1	131	1	PRO1_MALDO	Q9xf40 malus domes
35	5	16.1	131	1	PRO3_HEVBR	Q9m7n0 hevea bras
36	5	16.1	131	1	PRO5_HEVBR	Q9m7m8 hevea bras
37	5	16.1	131	1	PRO6_HEVBR	Q9le18 hevea bras
38	5	16.1	131	1	PROF_PYRGO	Q9xf38 pyrus commu
39	5	16.1	134	1	H34_MOUSE	P02301 mus musculu
40	5	16.1	134	1	PRO1_OLLEU	O24169 olea europ
41	5	16.1	134	1	PRO2_OLLEU	O24170 olea europ
42	5	16.1	134	1	PRO3_OLLEU	O24171 olea europ
43	5	16.1	150	1	PAD1_RHILO	Q9eyu1 rhizobium l
44	5	16.1	150	1	PAD2_RHILO	Q98290 rhizobium l
45	5	16.1	160	1	Y650_TREPA	O83656 treponema p
46	5	16.1	165	1	YC36_PORPU	P51273 porphyra pu
47	5	16.1	170	1	CT29_MOUSE	Q94742 mus musculu
48	5	16.1	171	1	ALL8_OLLEU	Q9m7r0 olea europ
49	5	16.1	173	1	NU6M_STRCA	O21407 struthio ca
50	5	16.1	173	1	RL10_SYNY3	P23350 synechocyst
51	5	16.1	175	1	CRAB_BOVIN	P02510 bos taurus
52	5	16.1	175	1	CRAB_HUMAN	P02511 homo sapien
53	5	16.1	175	1	CRAB_MESAU	P05811 mesocricetu
54	5	16.1	175	1	CRAB_MOUSE	P23927 mus musculu
55	5	16.1	175	1	CRAB_RABIT	P41316 oryctolagus
56	5	16.1	175	1	CRAB_RAT	P23928 rattus norv
57	5	16.1	175	1	Y773_METUA	Q58183 methanococc
58	5	16.1	181	1	RM06_ACACA	P46765 acanthamoeb
59	5	16.1	183	1	TFB_BPT4	P03739 bacterioph
60	5	16.1	192	1	COMK_BACSU	P40396 bacillus su
61	5	16.1	192	1	MOBA_PYRHO	O58708 pyrococcus
62	5	16.1	193	1	AX2D_PHAAR	O24542 phaseolus a
63	5	16.1	194	1	CSR3_HUMAN	P50461 homo sapien
64	5	16.1	194	1	CSR3_MOUSE	P50462 mus musculu
65	5	16.1	194	1	CSR3_RAT	P50463 rattus norv
66	5	16.1	198	1	DENR_HUMAN	Q43583 homo sapien
67	5	16.1	198	1	DENR_MOUSE	Q9cqj6 mus musculu
68	5	16.1	200	1	CT29_HUMAN	Q9mus5 homo sapien
69	5	16.1	206	1	RL3_BORBU	P94267 borrelia bu
70	5	16.1	210	1	YGEK_ECO57	P58664 escherichia
71	5	16.1	210	1	YGEK_ECOLI	Q48791 escherichia
72	5	16.1	218	1	GT27_SCHWA	P35661 schistosoma
73	5	16.1	220	1	FGF3_CHICK	P48801 gallus gall
74	5	16.1	224	1	SGN2_MOUSE	O55101 mus musculu
75	5	16.1	226	1	RB3C_BOVIN	P10949 bos taurus
76	5	16.1	229	1	LIPB_XYLFA	Q9pdv9 xyella fas
77	5	16.1	234	1	SGN2_RAT	O54980 rattus norv
78	5	16.1	235	1	TIRA_HUMAN	P58753 homo sapien
79	5	16.1	236	1	28XD_MYCLE	P19361 mycobacteri
80	5	16.1	237	1	60IM_PROMI	P22833 proteus mir
81	5	16.1	237	1	FGF3_XENLA	P36386 xenopus lae
82	5	16.1	239	1	FGF3_HUMAN	P11487 homo sapien
83	5	16.1	241	1	ISPD_MYCLE	Q9ccw6 mycobacteri
84	5	16.1	245	1	FGF3_MOUSE	P05524 mus musculu
85	5	16.1	246	1	DAPB_CHLMU	Q9pk30 chlamydia m
86	5	16.1	256	1	FGF3_BRARE	P48802 brachydanio
87	5	16.1	257	1	EUTC_RHOER	Q59782 rhodococcus
88	5	16.1	261	1	TFPB_MORBO	P20666 moraxella b
89	5	16.1	263	1	AMPM_SALTY	P10882 salmonella
90	5	16.1	267	1	BLAC_SERFO	P80345 serrattia fo
91	5	16.1	269	1	DAPF_CHLMU	Q9p1w2 chlamydia m
92	5	16.1	273	1	NPSN_DROME	Q9vzk0 drosophila
93	5	16.1	275	1	DAPF_CHLTR	Q84437 chlamydia t
94	5	16.1	286	1	HENK_VIBCH	Q9kq26 vibrio chol
95	5	16.1	291	1	KPPP_ALCEU	P19924 alcaligenes
96	5	16.1	291	1	KPPP_ALCEU	P19923 alcaligenes
97	5	16.1	294	1	YSHA_ECOLI	P25887 escherichia
98	5	16.1	295	1	RSP4_BOVIN	P26452 bos taurus
99	5	16.1	295	1	RSP4_CRIGR	P38982 cricetus
100	5	16.1	295	1	RSP4_HUMAN	P08865 homo sapien

```

CRAB COLLI          STANDARD;          PRT;          52 AA.
ID  CRAB COLLI          STANDARD;          PRT;          52 AA.
AC  Q12987;
DT  15-DEC-1998 (Rel. 37, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  15-DEC-1998 (Rel. 37, Last annotation update)
DE  Alpha crystallin B chain (Alpha(B)-crystallin) (Fragment).
GN  CRVAB.
OS  Columba livia (Domestic pigeon).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
OX  NCBI_TaxID=9932;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=97271643; PubMed=9126559;
RA  Caspers G.J., Uit de Weerd D., Wattel J., de Jong W.W.;
RT  "Alpha-crystallin sequences support a galliform/anseriform clade.";
RL  Mol. Phylogenet. Evol. 7:185-188(1997).
CC  -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
CC  OF THE LENS.
CC  -!- SUBUNIT: AGGREGATES WITH HOMOLOGOUS PROTEINS, INCLUDING ALPHA-A-
CC  CRYSTALLIN AND THE SMALL HEAT SHOCK PROTEIN HSP28, TO FORM LARGE
CC  HETEROMERIC COMPLEXES.
CC  -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC  FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; X96594; CAA65412.1; -
CC  InterPro; IPR003090; Crystallin_N.
CC  InterPro; IPR002068; Hsp20.
CC  Pfam; PF00525; crystallin; 1.
CC  ProDom; PD001193; Crystallin_N; 1.
CC  ProSITE; PS01031; HSP20; PARTIAL.
KW  Eye lens protein.
FT  NON_TER 1
FT  NON_TER 52
SQ  SEQUENCE 52 AA; 6078 MW; B4FC7ECF2BB79FE5 CRC64;

Query Match 19.4%; Score 6; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 IFGEHL 15
DB 19 IFGEHL 24

RESULT 2
CRAB TURME
ID  CRAB TURME          STANDARD;          PRT;          52 AA.
AC  O12995;
DT  15-DEC-1998 (Rel. 37, Created)
DT  15-DEC-1998 (Rel. 37, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Alpha crystallin B chain (Alpha(B)-crystallin) (Fragment).
GN  CRVAB.
OS  Turdus merula (Blackbird).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Passeriformes; Turdidae; Turdus.
OX  NCBI_TaxID=9187;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=97271643; PubMed=9126559;
RA  Caspers G.J., Uit de Weerd D., Wattel J., de Jong W.W.;
RT  "Alpha-crystallin sequences support a galliform/anseriform clade.";
RL  Mol. Phylogenet. Evol. 7:185-188(1997).

```

```

CC  -!- FUNCTION: MAY CONTRIBUTE TO THE TRANSPARENCY AND REFRACTIVE INDEX
CC  OF THE LENS.
CC  -!- SUBUNIT: AGGREGATES WITH HOMOLOGOUS PROTEINS, INCLUDING ALPHA-A-
CC  CRYSTALLIN AND THE SMALL HEAT SHOCK PROTEIN HSP28, TO FORM LARGE
CC  HETEROMERIC COMPLEXES.
CC  -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC  FAMILY.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
CC  EMBL; X96595; CAA65413.1; -
CC  InterPro; IPR003090; Crystallin_N.
CC  InterPro; IPR002068; Hsp20.
CC  Pfam; PF00525; crystallin; 1.
CC  ProDom; PD001193; Crystallin_N; 1.
CC  ProSITE; PS01031; HSP20; PARTIAL.
KW  Eye lens protein.
FT  NON_TER 1
FT  NON_TER 52
SQ  SEQUENCE 52 AA; 6143 MW; 82271ECE2619EF6F CRC64;

Query Match 19.4%; Score 6; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 IFGEHL 15
DB 19 IFGEHL 24

RESULT 3
Y039 METJA
ID  Y039 METJA          STANDARD;          PRT;          115 AA.
AC  Q60351;
DT  01-NOV-1997 (Rel. 35, Created)
DT  01-NOV-1997 (Rel. 35, Last sequence update)
DT  16-OCT-2001 (Rel. 40, Last annotation update)
DE  Hypothetical protein MJ0039.
GN  MJ0039.
OS  Methanococcus jannaschii.
OC  Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC  Methanocaldococcaceae; Methanocaldococcus.
OX  NCBI_TaxID=2190;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX  MEDLINE=96337999; PubMed=8688087;
RA  Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA  Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA  Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA  Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glöck A.,
RA  Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA  Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA  Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA  Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT  "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT  jannaschii.";
RL  Science 273:1058-1073(1996).
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 ; Search time 14.632 Seconds
(without alignments)
546.722 Million cell updates/sec

Title: US-10-045-180A-6
Perfect score: 31
Sequence: 1 ICHCRVLYCIRGEHLGGTCFILGRYPICCY 31

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 830525 seqs, 258052604 residues

Word size : 0

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : SPTREMBL.23.*

1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phase:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rviro:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	19.4	39	13	Q8J30
2	6	19.4	39	13	Q8J31
3	6	19.4	124	16	Q8KQ95
4	6	19.4	131	16	Q8UK07
5	6	19.4	136	2	O69489 mycobacteri
6	6	19.4	191	8	Q8WDD5 taenia soli
7	6	19.4	212	16	Q8XER3 salmoneilla
8	6	19.4	233	5	Q8MUQ6 anopheles g
9	6	19.4	236	13	Q90Z77
10	6	19.4	250	5	Q9VA75
11	6	19.4	268	16	Q97F29
12	6	19.4	287	16	Q9S15
13	6	19.4	294	2	Q46029
14	6	19.4	236	8	Q95813
15	6	19.4	296	8	Q8M675
16	6	19.4	296	8	Q8HN32

17	6	19.4	297	8	Q9B8W5
18	6	19.4	297	8	Q8M6K9
19	6	19.4	300	8	Q9B8Y4
20	6	19.4	300	8	Q34521
21	6	19.4	300	16	Q92AM6
22	6	19.4	300	16	Q8Y6J0
23	6	19.4	309	11	Q8VGP0
24	6	19.4	313	3	O59869
25	6	19.4	323	17	Q8TWC8
26	6	19.4	350	16	Q981Z0
27	6	19.4	352	16	Q9AA86
28	6	19.4	355	16	Q9RJU4
29	6	19.4	359	16	Q989G1
30	6	19.4	386	16	Q927W8
31	6	19.4	394	16	Q9K2D1
32	6	19.4	396	16	Q8EQ89
33	6	19.4	408	2	Q8KRQ4
34	6	19.4	408	16	O05898
35	6	19.4	410	6	Q8WMT0
36	6	19.4	410	16	Q9CCJ5
37	6	19.4	412	11	Q8BSQ6
38	6	19.4	419	16	Q8CJT2
39	6	19.4	423	11	Q8K1W9
40	6	19.4	423	17	Q8ZU85
41	6	19.4	424	11	Q8K1V2
42	6	19.4	424	11	Q8K1N9
43	6	19.4	424	11	Q8K1Q9
44	6	19.4	452	5	Q25631
45	6	19.4	452	11	Q9CUL7
46	6	19.4	459	5	Q9BI89
47	6	19.4	462	5	O01412
48	6	19.4	468	16	Q98C98
49	6	19.4	491	16	Q91ID6
50	6	19.4	506	16	Q8UF94
51	6	19.4	513	13	Q918V9
52	6	19.4	514	2	Q9FB48
53	6	19.4	518	16	Q32041
54	6	19.4	523	16	Q9EVP4
55	6	19.4	528	5	Q9VGT3
56	6	19.4	543	3	O94342
57	6	19.4	551	17	Q8TVW6
58	6	19.4	597	13	Q91871
59	6	19.4	599	16	Q8PPA1
60	6	19.4	599	16	Q8P3S5
61	6	19.4	639	3	O43113
62	6	19.4	694	13	Q91870
63	6	19.4	716	16	Q911V5
64	6	19.4	836	16	Q9PEF0
65	6	19.4	840	10	O64804
66	6	19.4	862	6	Q9BE52
67	6	19.4	873	2	Q9S0Z4
68	6	19.4	1018	1	Q8NKK1
69	6	19.4	1170	4	Q8NE11
70	6	19.4	1170	11	Q8CDL4
71	6	19.4	1170	11	Q8BND3
72	6	19.4	1219	4	Q9P2L0
73	6	19.4	2635	12	O40942
74	6	19.4	2635	12	P88955
75	5	16.1	46	9	Q38295
76	5	16.1	50	12	Q9QT23
77	5	16.1	52	16	Q8F683
78	5	16.1	54	10	Q8GYC8
79	5	16.1	57	16	Q92WF2
80	5	16.1	57	17	Q96X71
81	5	16.1	59	2	Q9KJ7
82	5	16.1	60	1	O57699
83	5	16.1	61	3	P89500
84	5	16.1	62	3	Q96V76
85	5	16.1	62	16	Q8FE45
86	5	16.1	63	9	Q8HAM2
87	5	16.1	64	11	Q91Y98
88	5	16.1	66	16	Q935L9
89	5	16.1	68	2	Q05625

Q9B8W5	taenia cras
Q8M6K9	taenia soli
Q9B8Y4	fasciola he
Q34521	fasciola he
Q92AM6	listeria in
Q8Y6J0	listeria mo
Q8VGP0	mus musculus
O59869	orpinomyces
Q8TWC8	methanopyru
Q981Z0	rhizobium l
Q9AA86	caulobacter
Q9RJU4	streptomyce
Q989G1	rhizobium l
Q927W8	chlamydia p
Q9K2D1	chlamydia p
Q8EQ89	oceanobacteri
Q8KRQ4	mycobacteri
O05898	mycobacteri
Q8WMT0	megaderma l
Q9CCJ5	mycobacteri
Q8BSQ6	mus musculus
Q8CJT2	streptomyce
Q8K1W9	bathyergus
Q8ZU85	pyrobaculum
Q8K1V2	echinys chr
Q8K1N9	thryonomys
Q8K1Q9	petromus ty
Q25631	onchocerca
Q9CUL7	mus musculus
Q9BI89	caenorhabdi
O01412	onchocerca
Q98C98	rhizobium l
Q91ID6	pseudomonas
Q8UF94	agrobacteri
Q918V9	raja eglant
Q9FB48	corynebacte
O32041	bacillus su
Q9EVP4	mycoplasma
Q9VGT3	drosophila
O94342	schizosacch
Q8TVW6	methanopyru
Q91871	xenopus lae
Q8PPA1	xanthomonas
Q8P3S5	xanthomonas
O43113	neurospora
Q91870	xenopus lae
Q911V5	pseudomonas
Q9PEF0	xylella fas
O64804	arabidopsis
Q9BE52	macaca fasc
Q9S0Z4	escherichia
Q8NKK1	acidianus a
Q8NE11	homo sapien
Q8CDL4	mus musculus
Q8BND3	mus musculus
Q9P2L0	homo sapien
O40942	kaposi's sa
P88955	lactococcus
Q38295	lactococcus
Q9QT23	tt virus. o
Q8F683	leptospira
Q8GYC8	arabidopsis
Q92WF2	rhizobium m
Q96X71	sulfolobus
Q9KJ7	hafnia alve
O57699	acidianus a
P89500	saccharomyc
Q96V76	calonelectria
Q8FE45	escherichia
Q8HAM2	burkholderi
Q91Y98	mus musculus
Q935L9	salmonella
Q05625	staphylococ

90 5 16.1 70 3 Q8X1M4
 91 5 16.1 71 3 Q8X1M2
 92 5 16.1 74 10 Q93VM1
 93 5 16.1 77 4 Q43207
 94 5 16.1 77 9 Q8H9Y3
 95 5 16.1 77 10 Q9S955
 96 5 16.1 77 12 Q9DHU7
 97 5 16.1 82 8 Q94UZ4
 98 5 16.1 83 8 Q8SKE8
 99 5 16.1 84 16 Q8Z177
 100 5 16.1 85 3 Q8X1M0

Q8X1M4 calonectria
 Q8X1M2 calonectria
 Q93VM1 oryza sativ
 Q43207 homo sapien
 Q8H9Y3 pseudomonas
 Q9S955 mesembryant
 Q9DHU7 yaba-like d
 Q94UZ4 gracula rel
 Q8SKE8 gracula rel
 Q8Z177 versinia pe
 Q8X1M0 calonectria

ALIGNMENTS

RESULT 1
 Q8JJ30
 ID Q8JJ30 PRELIMINARY; PRT; 39 AA.
 AC Q8JJ30;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Alpha-B-crystallin (Fragment).
 GN ABC.
 OS Ficedula albicollis.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Passeriformes; Muscicapidae; Ficedula.
 OX NCBI_TaxID=59894;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B8;
 RX MEDLINE=21918460; PubMed=11918793;
 RA Primer C.R., Borge T., Lindell J., Saetre G.-P.;
 RT "Single-nucleotide polymorphism characterization in species with
 RT limited available sequence information: high nucleotide diversity
 RT revealed in the avian genome.";
 RL Mol. Ecol. 11:603-612(2002).
 DR EMBL; AF454243; AAM22922.1; -;
 DR InterPro; IPR003090; Crystallin_N.
 DR Pfam; PF00525; Crystallin; 1.
 DR ProDom; PD001193; Crystallin_N; 1.
 FT NON_TER 1 1
 FT NON_TER 39 39
 SQ SEQUENCE 39 AA; 4442 MW; 24034FE4C5BD1E55 CRC64;
 Query Match 19.4%; Score 6; DB 13; Length 39;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 IFGEHL 15
 Db 12 IFGEHL 17

RESULT 2
 Q8JJ31
 ID Q8JJ31 PRELIMINARY; PRT; 39 AA.
 AC Q8JJ31;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Alpha-B-crystallin (Fragment).
 GN ABC.
 OS Ficedula hypoleuca.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Passeriformes; Muscicapidae; Ficedula.
 OX NCBI_TaxID=46689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Spl;
 RX MEDLINE=21918460; PubMed=11918793;
 RA Primer C.R., Borge T., Lindell J., Saetre G.-P.;

RT "Single-nucleotide polymorphism characterization in species with
 RT limited available sequence information: high nucleotide diversity
 RT revealed in the avian genome.";
 RL Mol. Ecol. 11:603-612(2002).
 DR EMBL; AF454242; AAM22921.1; -;
 DR InterPro; IPR003090; Crystallin_N.
 DR Pfam; PF00525; Crystallin; 1.
 DR ProDom; PD001193; Crystallin_N; 1.
 FT NON_TER 1 1
 FT NON_TER 39 39
 SQ SEQUENCE 39 AA; 4442 MW; 24034FE4C5BD1E55 CRC64;

Query Match 19.4%; Score 6; DB 13; Length 39;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 IFGEHL 15
 Db 12 IFGEHL 17

RESULT 3
 Q9KQ95
 ID Q9KQ95 PRELIMINARY; PRT; 124 AA.
 AC Q9KQ95;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical protein VC2105.
 GN VC2105.
 OS Vibrio cholerae.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Vibrio.
 OX NCBI_TaxID=666;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=El Tor N16961 / Serotype O1;
 RX MEDLINE=20406833; PubMed=10952301;
 RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
 RA Emolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
 RT cholerae.";
 RL Nature 406:477-483(2000).
 DR EMBL; AB004284; AAF95251.1; -;
 DR TIGR; VC2105; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 124 AA; 14590 MW; 6E68B678976CBBF3 CRC64;

Query Match 19.4%; Score 6; DB 16; Length 124;
 Best Local Similarity 100.0%; Pred. No. 44;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 21 ILGERY 26
 Db 78 ILGERY 83

RESULT 4
 Q8UK07
 ID Q8UK07 PRELIMINARY; PRT; 131 AA.
 AC Q8UK07;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Hypothetical protein Atu5317.
 GN ATU5317.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OG Plasmid AT.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 ; Search time 18.352 Seconds
(without alignments)
268.119 Million cell updates/sec

Title: US-10-045-180A-6
Perfect score: 31
Sequence: 1 ICHCRVLYCIGFHLGGTCFLIGERYPICCY 31

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1107863 seqs, 159726573 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database : A Geneseq 19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	31	20	AAV01606
2	31	100.0	94	20	AAV01603
3	31	100.0	94	20	AAV01604
4	6	19.4	37	23	ABP26644
5	6	19.4	64	22	AAU58954
6	6	19.4	81	21	AAU59901
7	6	19.4	81	21	AAU59901
8	6	19.4	111	22	AAU57814
9	6	19.4	137	23	AAU05700
					Fragment of human
					Amino acid sequenc
					Amino acid sequenc
					Streptococcus poly
					Propionibacterium
					Arabidopsis thalia
					Arabidopsis thalia
					Propionibacterium
					M. tuberculosis an

10	6	19.4	179	22	ABG52130	Human liver peptid
11	6	19.4	179	22	ABB32059	Peptide #4710 enco
12	6	19.4	179	22	ABB37307	Peptide #4813 enco
13	6	19.4	179	22	ABB22600	Protein #4599 enco
14	6	19.4	179	22	AAU58003	Human brain expres
15	6	19.4	179	22	AAU70444	Human bone marrow
16	6	19.4	179	22	AAU18276	Peptide #4710 enco
17	6	19.4	179	22	AAU30765	Peptide #4802 enco
18	6	19.4	179	22	AAU05888	Peptide #4570 enco
19	6	19.4	179	23	ABG40078	Human peptid enco
20	6	19.4	249	23	AAU91178	Pantothenate kinas
21	6	19.4	275	22	ABB70744	Drosophila melanog
22	6	19.4	294	22	ABG06705	Novel human diagn
23	6	19.4	296	22	AAU35611	Haemophilus influe
24	6	19.4	300	23	ABB49371	Listeria monocytog
25	6	19.4	337	20	AAU35212	Chlamydia pneumoni
26	6	19.4	368	23	ABB04596	Lettuce big-vein v
27	6	19.4	408	22	AAU81217	Mycobacterium tube
28	6	19.4	456	22	AAU88530	Haemophilus influe
29	6	19.4	456	23	AAU91461	Haemophilus influe
30	6	19.4	498	22	ABG24816	Novel human diagn
31	6	19.4	499	22	ABG24685	Novel human diagn
32	6	19.4	512	24	ABJ18765	Pseudomonas aerugi
33	6	19.4	528	22	ABB67113	Drosophila melanog
34	6	19.4	545	23	AAU75783	Human protein phos
35	6	19.4	716	24	ABJ18808	Pseudomonas aerugi
36	6	19.4	4655	17	AAU97208	Human calcium sens
37	6	19.4	4655	17	AAU97209	Human placental ca
38	6	19.4	4655	17	AAU97210	Human kidney calci
39	6	19.4	4655	17	AAU97211	Human parathyroid
40	6	19.4	4655	19	AAU43311	Human calcium sens
41	6	19.4	4655	19	AAU43312	Human placental ca
42	6	19.4	4655	19	AAU43313	Human kidney calci
43	6	19.4	4655	19	AAU43314	Human parathyroid
44	6	19.4	4655	24	ABP56836	Human megalin prot
45	6	19.4	4689	22	ABG04530	Novel human diagn
46	5	16.1	13	15	AAU46829	pH 2.5 acid phosph
47	5	16.1	19	23	ABP46151	Human Blys binding
48	5	16.1	20	22	ABG65927	Human secreted pro
49	5	16.1	21	23	AAU89614	Insulin/insulin-li
50	5	16.1	22	20	AAU01451	Secreted protein e
51	5	16.1	24	20	AAU12811	Human 5' EST secre
52	5	16.1	24	22	ABP35824	LHC_XENLA positio
53	5	16.1	27	20	AAU27707	Human secreted pro
54	5	16.1	40	21	AAU54168	Human pancreatic c
55	5	16.1	46	22	ABP35780	Dcdc2 positions p
56	5	16.1	52	22	ABU03954	Human musculoskele
57	5	16.1	52	23	ABP34118	Human ORF3091 prot
58	5	16.1	52	23	ABP34254	Human glycoprotein
59	5	16.1	52	24	ABU13248	Novel human muscul
60	5	16.1	53	23	ABP31305	Human ORF278 prote
61	5	16.1	55	23	ABP32567	Human ORF1540 prot
62	5	16.1	55	23	ABP01122	Human ORFX protein
63	5	16.1	56	22	ABG52884	Human liver peptid
64	5	16.1	56	22	AAU59558	Propionibacterium
65	5	16.1	56	22	AAU66153	Propionibacterium
66	5	16.1	56	22	ABG02704	Novel human diagn
67	5	16.1	56	22	ABG38036	Peptide #5542 enco
68	5	16.1	56	22	AAU58666	Human brain expres
69	5	16.1	56	22	AAU71172	Human bone marrow
70	5	16.1	56	22	AAU31453	Peptide #5490 enco
71	5	16.1	56	23	ABG40969	Human peptid enco
72	5	16.1	59	22	AAU88038	Human immune/haema
73	5	16.1	65	22	ABU16086	Human nervous syst
74	5	16.1	66	21	AAU57014	Human prostate can
75	5	16.1	67	22	AAU45709	Propionibacterium
76	5	16.1	67	23	ABG3264	Human prostate spe
77	5	16.1	68	23	ABP03828	Human ORFX protein
78	5	16.1	69	23	ABP42696	Human ovarian anti
79	5	16.1	70	22	AAU84241	Human immune/haema
80	5	16.1	71	22	ABU69664	Drosophila melanog
81	5	16.1	71	22	AAU49208	Propionibacterium
82	5	16.1	71	23	ABP05221	Human ORFX protein

83 5 16.1 72 22 AAU52524
 84 5 16.1 72 22 AAU57734
 85 5 16.1 73 22 AAU40466
 86 5 16.1 73 22 AM82541
 87 5 16.1 73 23 ABP04083
 88 5 16.1 75 22 AAU50025
 89 5 16.1 75 22 ABB16158
 90 5 16.1 77 23 ABP35207
 91 5 16.1 78 22 AAU46902
 92 5 16.1 78 22 AM80786
 93 5 16.1 78 22 AM81231
 94 5 16.1 78 22 AM81596
 95 5 16.1 78 22 AM81841
 96 5 16.1 78 24 ABU04778
 97 5 16.1 78 24 ABU04779
 98 5 16.1 78 24 ABU04780
 99 5 16.1 78 24 ABU04781
 100 5 16.1 79 22 AAU46468

ALIGNMENTS

RESULT 1

AAU01606
 ID AAU01606 standard; peptide; 31 AA.

XX
 AC AAU01606;

XX
 DT 18-JUN-1999 (first entry)

XX
 DE Fragment of human defensin (Def-X) protein.

XX
 KW Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
 KW cytostatic; anticancer; inflammation; tissue repair;
 KW endocrine regulation; corticostatic regulation; cancer; melanoma;
 KW AIDS; immune deficiency; psoriasis.

XX
 OS Homo sapiens.

XX
 PN FR2767832-A1.

XX
 PD 05-MAR-1999.

XX
 PF 29-AUG-1997; 97FR-0010823.

XX
 PR 29-AUG-1997; 97FR-0010823.

XX
 PA (GEST) GENSET SA.

XX
 PI Bougueleret L, Chumakov I;

XX
 DR WPI; 1999-183266/16.

XX
 PT Human defensin polypeptide Def-X - useful as antimicrobial agent,
 PT anticancer agent, pesticide, etc.

XX
 PS Claim 2; Page 33; 56pp; French.

XX
 CC The present sequence represents human defensin (Def-X) fragment. The
 CC Def-X polypeptide can be used as an antimicrobial, antiparasitic agent
 CC or a pesticide. The Def-X polypeptide can be used as a cytostatic
 CC (especially anticancer) agent, and as an agent for modulating processes
 CC of inflammation, tissue repair and endocrine (especially corticostatic)
 CC regulation. The polypeptide can be used in a composition for external
 CC topical use, especially in a cosmetic composition. Compositions
 CC containing Def-X can be used for prevention and treatment of microbial
 CC and parasitic infections, especially where the microbial or parasitic
 CC infections are Gram-positive or -negative bacterial infections or
 CC mycobacterial, fungal or spirochaete infections, or where the viral
 CC infections are associated with enveloped viruses, especially HSV and HIV.
 CC The compositions can be used for prevention and/or treatment of cancers,
 CC especially melanomas, or liver cancer, prostate cancer, non-small-cell

CC lung cancer or colorectal carcinoma, and for enhancing immunity,
 CC especially in the case of AIDS, or preventing immune deficiency,
 CC especially in the treatment of psoriasis, as well as for modulating
 CC inflammatory processes, especially in the case of chronic inflammatory
 CC disorders.

XX
 SQ Sequence 31 AA;

Query Match 100.0%; Score 31; DB 20; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.1e-24;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ICHCRVLYCIFGSHLGGTCFILGERYPICCY 31

Db 1 ICHCRVLYCIFGSHLGGTCFILGERYPICCY 31

RESULT 2

AAU01603
 ID AAU01603 standard; protein; 94 AA.

XX
 AC AAU01603;

XX
 DT 18-JUN-1999 (first entry)

XX
 DE Amino acid sequence of the human defensin (Def-X) protein.

XX
 KW Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
 KW cytostatic; anticancer; inflammation; tissue repair;
 KW endocrine regulation; corticostatic regulation; cancer; melanoma;
 KW AIDS; immune deficiency; psoriasis.

XX
 OS Homo sapiens.

XX
 PN FR2767832-A1.

XX
 PD 05-MAR-1999.

XX
 PF 29-AUG-1997; 97FR-0010823.

XX
 PR 29-AUG-1997; 97FR-0010823.

XX
 PA (GEST) GENSET SA.

XX
 PI Bougueleret L, Chumakov I;

XX
 DR WPI; 1999-183266/16.

XX
 PT Human defensin polypeptide Def-X - useful as antimicrobial agent,
 PT anticancer agent, pesticide, etc.

XX
 PS Claim 1; Fig 1; 56pp; French.

XX
 CC The present sequence represents human defensin (Def-X). The Def-X
 CC polypeptide can be used as an antimicrobial, antiparasitic agent or
 CC a pesticide. The Def-X polypeptide can be used as a cytostatic
 CC (especially anticancer) agent, and as an agent for modulating processes
 CC of inflammation, tissue repair and endocrine (especially corticostatic)
 CC regulation. The polypeptide can be used in a composition for external
 CC topical use, especially in a cosmetic composition. Compositions
 CC containing Def-X can be used for prevention and treatment of microbial
 CC and parasitic infections, especially where the microbial or parasitic
 CC infections are Gram-positive or -negative bacterial infections or
 CC mycobacterial, fungal or spirochaete infections, or where the viral
 CC infections are associated with enveloped viruses, especially HSV and HIV.
 CC The compositions can be used for prevention and/or treatment of cancers,
 CC especially melanomas, or liver cancer, prostate cancer, non-small-cell
 CC lung cancer or colorectal carcinoma, and for enhancing immunity,
 CC especially in the case of AIDS, or preventing immune deficiency,
 CC especially in the treatment of psoriasis, as well as for modulating
 CC inflammatory processes, especially in the case of chronic inflammatory
 CC disorders.

```
SQ Sequence 94 AA;
  Query Match 100.0%; Score 31; DB 20; Length 94;
  Best Local Similarity 100.0%; Pred. No. 3e-24; Mismatches 0; Indels 0; Gaps 0;
  Matches 31; Conservative 0;

QY 1 ICHCRVLYCIFGEHLGGTCFILGERYPICCY 31
   |||||
DB 64 ICHCRVLYCIFGEHLGGTCFILGERYPICCY 94

RESULT 3
AAY01604
ID AAY01604 standard; protein; 94 AA.
XX
AC AAY01604;
XX
DT 18-JUN-1999 (first entry)
XX
DE Amino acid sequence of the human defensin (Def-X) protein.
XX
KW Human defensin; Def-X; antimicrobial; antiparasitic; pesticide;
KW cytostatic; anticancer; inflammation; tissue repair;
KW endocrine regulation; corticostatic regulation; cancer; melanoma;
KW AIDS; immune deficiency; psoriasis.
XX
OS Homo sapiens.
XX
FN FR2767832-A1.
XX
PD 05-MAR-1999.
XX
PF 29-AUG-1997; 97FR-0010823.
XX
PR 29-AUG-1997; 97FR-0010823.
XX
PA (GEST ) GENSET SA.
XX
PI Bougueleret L, Chumakov I;
XX
DR WPI; 1999-183266/16.
DR N-PSDB; AAX26697.
XX
PT Human defensin polypeptide Def-X - useful as antimicrobial agent,
PT anticancer agent, pesticide, etc.
XX
PS Disclosure; Fig 3; 56pp; French.
XX
CC The present sequence represents human defensin (Def-X). The Def-X
CC polypeptide can be used as an antimicrobial, antiparasitic agent or
CC a pesticide. The Def-X polypeptide can be used as a cytostatic
CC (especially anticancer) agent, and as an agent for modulating processes
CC of inflammation, tissue repair and endocrine (especially corticostatic)
CC regulation. The polypeptide can be used in a composition for external
CC topical use, especially in a cosmetic composition. Compositions
CC containing Def-X can be used for prevention and treatment of microbial
CC and parasitic infections, especially where the microbial or parasitic
CC infections are Gram-positive or -negative bacterial infections or
CC mycobacterial, fungal or spirochaete infections, or where the viral
CC infections are associated with enveloped viruses, especially HSV and HIV.
CC The compositions can be used for prevention and/or treatment of cancers,
CC especially melanomas, or liver cancer, prostate cancer, non-small-cell
CC lung cancer or colorectal carcinoma, and for enhancing immunity,
CC especially in the case of AIDS, or preventing immune deficiency,
CC especially in the treatment of psoriasis, as well as for modulating
CC inflammatory processes, especially in the case of chronic inflammatory
CC disorders.
XX
SQ Sequence 94 AA;
  Query Match 100.0%; Score 31; DB 20; Length 94;
  Best Local Similarity 100.0%; Pred. No. 3e-24; Mismatches 0; Indels 0; Gaps 0;
  Matches 31; Conservative 0;

QY 1 ICHCRVLYCIFGEHLGGTCFILGERYPICCY 31
   |||||
DB 64 ICHCRVLYCIFGEHLGGTCFILGERYPICCY 94

SQ Sequence 94 AA;
  Query Match 19.4%; Score 6; DB 23; Length 37;
  Best Local Similarity 100.0%; Pred. No. 39;
  Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  Matches 21 ILGERY 26.
   |||||
```

```
QY 1 ICHCRVLYCIFGEHLGGTCFILGERYPICCY 31
   |||||
DB 64 ICHCRVLYCIFGEHLGGTCFILGERYPICCY 94

RESULT 4
ABP26644
ID ABP26644 standard; Protein; 37 AA.
XX
AC ABP26644;
XX
DT 02-JUL-2002 (first entry)
XX
DE Streptococcus polypeptide SEQ ID NO 2464.
XX
KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus agalactiae.
XX
FN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
PF 29-OCT-2001; 2001WO-GB04789.
XX
PR 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
PA (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
PI Telford J, Massignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
DR WPI; 2002-352536/38.
DR N-PSDB; ABN67275.
XX
PT New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 3396; 4525pp; English.
XX
CC The invention relates to a protein (ABP25413-ABP30895) from group B
CC streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (SI), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC (I). A composition comprising (I) or a nucleic acid encoding (I), may be
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 37 AA;
```

```
Db          7 ILGCRY 12
RESULT 5
AAU58954
ID AAU58954 standard; Protein; 64 AA.
XX
AC AAU58954;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein #19850.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US12865.
XX
PR 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
XX Skeiky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI: 2001-616774/71.
DR N-PSDB; AAS59596.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID No 20149; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 64 AA;
Query Match 19.4%; Score 6; DB 22; Length 64;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 14 HLGTC 19
|||||
```

```
Db          44 HLGTC 49
RESULT 6
AAG59901
ID AAG59901 standard; Protein; 81 AA.
XX
AC AAG59901;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 77532.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 23-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
```


89 5 16.1 235 10 US-09-860-696-2 Sequence 2, Appli
90 5 16.1 235 15 US-10-188-947-3 Sequence 3, Appli
91 5 16.1 236 16 US-10-080-170-21 Sequence 21, Appl
92 5 16.1 238 15 US-10-156-761-13747 Sequence 13747, A
93 5 16.1 239 9 US-09-822-485-6 Sequence 6, Appli
94 5 16.1 239 9 US-09-750-963-8 Sequence 8, Appli
95 5 16.1 239 9 US-09-902-773A-9 Sequence 9, Appli
96 5 16.1 239 9 US-09-425-021-11 Sequence 11, Appli
97 5 16.1 239 10 US-09-995-598-62 Sequence 62, Appli
98 5 16.1 239 11 US-09-345-373-21 Sequence 21, Appli
99 5 16.1 239 12 US-10-189-360-12 Sequence 12, Appli
100 5 16.1 239 12 US-10-374-207-6 Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-10-013-770-6
; Sequence 6, Application US/10013770
; Publication No. US20020115151A1
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/013,770
; FILING DATE: 10-Dec-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/486,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: MATURE PEPTIDE
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-013-770-6
Query Match 100.0%; Score 31; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.3e-25;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ICHCRVLYCIFGHELGTCFILGERYPICCY 31
Db 1 ICHCRVLYCIFGHELGTCFILGERYPICCY 31

RESULT 2
US-10-045-180A-6
; Sequence 6, Application US/10045180A

; Publication No. US20020182703A1
; GENERAL INFORMATION:
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilva
; TITLE OF INVENTION: Human Defensin Polypeptide Def-X, Genomic DNA and cDNA, Composit
; TITLE OF INVENTION: Containing them and Applications to Diagnosis and to Therapeuti
; FILE REFERENCE: GEN-100D1
; CURRENT APPLICATION NUMBER: US/10/045,180A
; CURRENT FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: US 09/486,580
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: PCT/FR98/01864
; PRIOR FILING DATE: 1998-08-28
; PRIOR APPLICATION NUMBER: FR 97/10823
; PRIOR FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 31
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(31)
; OTHER INFORMATION: Def-X mature peptide
US-10-045-180A-6
Query Match 100.0%; Score 31; DB 14; Length 31;
Best Local Similarity 100.0%; Pred. No. 2.3e-25;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ICHCRVLYCIFGHELGTCFILGERYPICCY 31
Db 1 ICHCRVLYCIFGHELGTCFILGERYPICCY 31
RESULT 3
US-10-013-770-3
; Sequence 3, Application US/10013770
; Publication No. US20020115151A1
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/013,770
; FILING DATE: 10-Dec-2001
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/486,580
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 94 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 17, 2003, 14:50:14 ; Search time 6.944 Seconds
(without alignments)
188.888 Million cell updates/sec

Title: US-10-045-180A-6
Perfect score: 31
Sequence: 1 ICHCRVLYCIFGEHLGGTCFLIGERYPICCY 31

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 328717 seqs, 42310858 residues

Word size : 0

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	31	4	US-09-486-580A-6
2	31	100.0	94	4	US-09-486-580A-3
3	6	19.4	111	4	US-09-107-532A-3707
4	6	19.4	136	4	US-09-252-991A-22070
5	6	19.4	147	4	US-09-252-991A-31557
6	6	19.4	323	4	US-09-252-991A-24673
7	6	19.4	337	4	US-09-198-452A-630
8	6	19.4	417	4	US-09-252-991A-28991
9	6	19.4	429	4	US-09-252-991A-20363
10	6	19.4	443	4	US-09-252-991A-31911
11	6	19.4	530	4	US-09-252-991A-18090
12	6	19.4	535	4	US-09-252-991A-28410
13	6	19.4	581	4	US-09-252-991A-25414
14	6	19.4	652	4	US-09-252-991A-30238
15	6	19.4	4654	3	US-08-476-515A-84
16	6	19.4	4655	3	US-08-652-877-84
17	6	19.4	4655	3	US-08-652-877-86
18	6	19.4	4655	3	US-08-652-877-88
19	6	19.4	4655	3	US-08-652-877-90
20	5	16.1	13	1	US-07-923-724-57
21	5	16.1	13	2	US-08-609-428A-57
22	5	16.1	13	2	US-08-374-652C-44
23	5	16.1	18	2	US-08-374-652C-84
24	5	16.1	18	2	US-08-374-652C-87
25	5	16.1	18	2	US-08-374-652C-88
26	5	16.1	70	3	US-08-851-843A-190
27	5	16.1	70	3	US-08-974-549A-309

28	5	16.1	70	3	US-08-854-050-190	Sequence 190, App
29	5	16.1	70	4	US-09-430-323-190	Sequence 190, App
30	5	16.1	75	6	5175383-2	Patent No. 5175383
31	5	16.1	95	1	US-08-455-896-7	Sequence 7, Appli
32	5	16.1	95	2	US-08-933-149-7	Sequence 7, Appli
33	5	16.1	95	2	US-09-082-343-7	Sequence 7, Appli
34	5	16.1	95	3	US-09-082-253-7	Sequence 7, Appli
35	5	16.1	95	3	US-08-821-451A-27	Sequence 27, Appl
36	5	16.1	95	3	US-09-263-810-27	Sequence 27, Appl
37	5	16.1	95	4	US-09-583-169-27	Sequence 27, Appl
38	5	16.1	95	4	US-09-162-622-7	Sequence 7, Appli
39	5	16.1	95	5	PCT-US96-08235-7	Sequence 7, Appli
40	5	16.1	112	4	US-08-311-731A-22	Sequence 22, Appl
41	5	16.1	115	4	US-09-252-991A-20440	Sequence 20440, A
42	5	16.1	149	4	US-09-252-991A-21985	Sequence 21985, A
43	5	16.1	157	4	US-09-252-991A-22237	Sequence 22237, A
44	5	16.1	159	1	US-08-597-236-9	Sequence 9, Appli
45	5	16.1	159	1	US-08-746-682A-9	Sequence 9, Appli
46	5	16.1	173	4	US-09-732-210-538	Sequence 538, App
47	5	16.1	183	1	US-08-441-629-15	Sequence 15, Appl
48	5	16.1	183	3	US-08-776-207-15	Sequence 15, Appl
49	5	16.1	183	4	US-09-507-773-15	Sequence 15, Appl
50	5	16.1	183	5	PCT-US95-09172-15	Sequence 15, Appl
51	5	16.1	190	1	US-08-441-629-4	Sequence 4, Appli
52	5	16.1	190	3	US-08-776-207-4	Sequence 4, Appli
53	5	16.1	190	4	US-09-507-773-4	Sequence 4, Appli
54	5	16.1	190	5	PCT-US95-09172-4	Sequence 4, Appli
55	5	16.1	193	4	US-09-252-991A-22938	Sequence 22938, A
56	5	16.1	194	1	US-08-616-368A-11	Sequence 11, Appl
57	5	16.1	194	2	US-08-739-485-8	Sequence 8, Appli
58	5	16.1	194	3	US-09-054-298-11	Sequence 11, Appl
59	5	16.1	194	3	US-08-818-655-11	Sequence 11, Appl
60	61	16.1	194	4	US-09-305-839-11	Sequence 11, Appl
61	5	16.1	197	4	US-09-198-452A-271	Sequence 271, App
62	5	16.1	205	4	US-09-107-532A-5214	Sequence 5214, Ap
63	5	16.1	225	4	US-09-107-532A-7302	Sequence 7302, Ap
64	5	16.1	228	4	US-09-252-991A-18371	Sequence 18371, A
65	5	16.1	228	4	US-09-252-991A-28331	Sequence 28331, A
66	5	16.1	231	4	US-09-252-991A-31713	Sequence 31713, A
67	5	16.1	232	4	US-08-733-622C-31	Sequence 31, Appl
68	5	16.1	233	4	US-09-328-352-7602	Sequence 7602, Ap
69	5	16.1	236	3	US-08-711-417C-174	Sequence 174, App
70	5	16.1	236	3	US-08-711-417C-175	Sequence 175, App
71	5	16.1	236	3	US-08-793-701-56	Sequence 56, Appl
72	5	16.1	236	3	US-08-793-701-58	Sequence 58, Appl
73	5	16.1	236	4	US-09-579-264-56	Sequence 56, Appl
74	5	16.1	237	4	US-09-579-264-58	Sequence 58, Appl
75	5	16.1	237	4	US-09-107-532A-6583	Sequence 6583, Ap
76	5	16.1	238	3	US-08-711-417C-176	Sequence 176, App
77	5	16.1	238	3	US-08-711-417C-177	Sequence 177, App
78	5	16.1	239	1	US-08-464-590A-11	Sequence 11, Appl
79	5	16.1	239	1	US-08-462-169B-11	Sequence 11, Appl
80	5	16.1	239	2	US-08-207-412B-14	Sequence 14, Appl
81	5	16.1	239	2	US-08-951-822-35	Sequence 35, Appl
82	5	16.1	239	3	US-09-103-079-11	Sequence 11, Appl
83	5	16.1	239	3	US-08-718-904-12	Sequence 12, Appl
84	5	16.1	239	3	US-09-023-082A-21	Sequence 21, Appl
85	5	16.1	239	3	US-09-093-585-11	Sequence 11, Appl
86	5	16.1	239	4	US-09-368-951-35	Sequence 35, Appl
87	5	16.1	239	4	US-09-425-021-11	Sequence 11, Appl
88	5	16.1	239	4	US-09-449-249-12	Sequence 12, Appl
89	5	16.1	239	4	US-09-390-207-18	Sequence 18, Appl
90	5	16.1	239	4	US-09-229-947-35	Sequence 35, Appl
91	5	16.1	240	4	US-09-417-721-7	Sequence 7, Appli
92	5	16.1	241	6	5175383-7	Patent No. 5175383
93	5	16.1	245	1	US-08-439-725A-11	Sequence 11, Appl
94	5	16.1	245	2	US-08-867-471-11	Sequence 11, Appl
95	5	16.1	245	2	US-08-438-439C-7	Sequence 7, Appli
96	5	16.1	245	2	US-08-438-439C-17	Sequence 17, Appl
97	5	16.1	245	3	US-08-705-245-12	Sequence 12, Appl
98	5	16.1	245	4	US-09-390-207-28	Sequence 28, Appl
99	5	16.1	246	4	US-09-252-991A-20602	Sequence 20602, A
100	5	16.1	246	4	US-09-252-991A-22967	Sequence 22967, A

ALIGNMENTS

```
RESULT 1
US-09-486-580A-6
; Sequence 6, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: MATURE PEPTIDE
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-09-486-580A-6

Query Match 100.0%; Score 31; DB 4; Length 31;
Best Local Similarity 100.0%; Pred. No. 1-2e-24;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ICHCRVLYCIFGHLGGTCFILGERYPICCY 31
Db 1 ICHCRVLYCIFGHLGGTCFILGERYPICCY 31

RESULT 2
US-09-486-580A-3
; Sequence 3, Application US/09486580A
; Patent No. 6329340
; GENERAL INFORMATION:
; APPLICANT: GENSET SA
; TITLE OF INVENTION: HUMAN DEFENSIN DEF-X GENE AND DNAC
; TITLE OF INVENTION: COMPOSITION CONTAINING SAME AND DIAGNOSTIC AND
; TITLE OF INVENTION: THERAPEUTIC APPLICATIONS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 550 West C Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/486,580A
; FILING DATE: FEBRUARY 25, 2000
; ATTORNEY/AGENT INFORMATION:
; NAME: Hart, Daniel
; REGISTRATION NUMBER: 40,637
; REFERENCE/DOCKET NUMBER: GENSET.064C1
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 94 AMINO ACIDS
; TYPE: AMINO ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL PEPTIDE
; LOCATION: 1..19
; FEATURE:
; NAME/KEY: PRO REGION
; LOCATION: 20..63
; FEATURE:
; NAME/KEY: MATURE PEPTIDE
; LOCATION: 64..94
; US-09-486-580A-3

Query Match 100.0%; Score 31; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 3e-24;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ICHCRVLYCIFGHLGGTCFILGERYPICCY 31
Db 64 ICHCRVLYCIFGHLGGTCFILGERYPICCY 94

RESULT 3
US-09-107-532A-5707
; Sequence 5707, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
```